

UNIVERSITY OF TARTU  
FACULTY OF SCIENCE AND TECHNOLOGY  
INSTITUTE OF MOLECULAR AND CELL BIOLOGY  
CHAIR OF MICROBIOLOGY AND VIROLOGY

**Generation of human papillomavirus type 11 and 18 genomes encoding hemagglutinin  
epitope-tagged E1 proteins**

Bachelor of Science Thesis

12 EAP

Viktor Levin

Supervisor: PhD Alla Piirsoo

TARTU 2018

## **Generation of human papillomavirus type 11 and 18 genomes encoding hemagglutinin epitope-tagged E1 proteins**

Human papillomaviruses (HPVs) are associated with numerous disorders ranging from warts to cancerous diseases. Although vaccination against several types of HPVs is available, a number of HPV-associated cancers remains stable. Therefore, studies on HPV are required for developing novel HPV-targeting drugs. Encoded by HPV E1 helicase is absolutely required for HPV replication. Therefore, it is a good target for developing antiviral drugs.

This project aimed at generating HPV11 and HPV18 genomes encoding human influenza hemagglutinin (HA)-tagged E1 proteins with the further goal to test their replication ability in U2OS cell line. The study revealed that HA-tagged E1 proteins maintained the replication of the modified HPV genomes at levels similar to wild-type HPVs (wtHPVs). Thus, this study opens possibilities to investigate HPV11 and HPV18 endogenous E1 proteins using highly specific antibodies against HA epitope and introduces a less expensive and universal method to study E1-associated processes in the HPV life cycle.

**Keywords:** E1 protein, HA-tag, human papillomavirus (HPV), replication

**CERCS code:** B230 Microbiology, bacteriology, virology, mycology

## **Inimese papilloomiviiruse tüüp 11 ja 18 hemagglutiniin märgisega E1 lugemisraami kodeerivate genoomide kloneerimine ja analüüs**

Inimese papillomaviirused (HPV) on seotud erinevate terviseprobleemidega alates tüükadest ja lõpetades pahaloomuliste kasvajatega. Kuigi vaktsineerimine mõne HPV vastu on juba olemas, HPV-dega seotud kasvajate hulk oluliselt ei vähene. Seetõttu on vajalik arendada uudseid HPV-de vastu ravimeid. Kodeeritud HPV genoomis E1 helikaas on vajalik HPV replikatsiooniks. Seetõttu, E1 on hea sihtmärk viirusevastaste ravimite arendamiseks.

Selle uuringu eesmärk oli luua HPV11 ja HPV18 genoomid, mis kodeerivad E1 valgu koos inimese gripi viiruse hemagglutiniin (HA)-märgisega ning nende replikatsiooni võimet kontrollimine U2OS rakku liinis. Uuringu tulemused näitasid, et muteeritud HPV genoomid replitseerusid sarnaselt metsik-tüüpi HPV-dega. Seega, see uuring avab võimalusi uurida HPV11 ja HPV18 endogeense E1 valguga seotud protsesse kasutades spetsiifilisi antikehi HA-märgise vastu.

**Märksõnad:** E1 valk, inimese papillomaviirus (HPV), HA-märgis, replikatsioon

**CERCS kood:** B230 Mikrobioloogia, bakterioloogia, viroloogia, mükoloogia

# CONTENT

CONTENT .....	3
LIST OF ABBREVIATIONS .....	5
INTRODUCTION .....	7
1. LITERATURE OVERVIEW .....	8
1.1 Papillomaviridae .....	8
1.2 Human Papillomaviruses .....	8
1.2.1 Cutaneous HPVs .....	9
1.2.2 Mucosal HPVs .....	10
1.2.3 Prevention or treatment of HPV infection .....	10
1.3 HPV infection cycle .....	11
1.4 HPV molecular biology .....	12
1.4.1 HPV genomic organization .....	12
1.4.2 Overview of viral proteins .....	13
1.5 E1 protein structure .....	16
1.6 Role of E1 in the life cycle of HPVs .....	17
1.6.1 E1 and replication of the HPV genome .....	18
1.6.2 Interactions between E1 and host cellular factors .....	19
1.6.3 E1 toxicity for the host cell and regulation of E1 .....	20
1.6.4 E1-induced DNA damage response .....	20
1.7 Difficulties to study the HPV life cycle .....	21
2. EXPERIMENTAL PART .....	22
2.1 Aims of the study .....	22
2.2 Materials and methods .....	23
2.2.1 Plasmids used in the present study .....	23

2.2.2 Generation of E1-HA HPV11/HPV18-pMC.BESPX constructs .....	23
2.2.3 Electroporation of the U2OS cell line and isolation of total DNA.....	28
2.2.4 Southern Blot.....	29
2.3 Results .....	31
2.3.1 Synthesis of fragments of HPV11/HPV18 DNA with the HA-tag encoding sequence.....	31
2.3.2 Cloning the PCR products into the vector and controlling them with restriction and DNA sequencing .....	32
2.3.3 Generation of E1-HA HPV11/HPV18 minicircle constructs .....	34
2.3.4 Comparison of the replication efficiencies of wtHPV11 with E1-HA HPV11 and wtHPV18 with E1-HA HPV18 minicircles.....	36
2.4 Discussion.....	39
CONCLUSION .....	40
RESÜMEE .....	41
REFERENCES .....	42
USED WEBSITES .....	49

# LIST OF ABBREVIATIONS

ara-BAD – *Escherichia coli* operon that encodes proteins required for the metabolism of L-arabinose under control of P<sub>BAD</sub> L-arabinose inducible promotor

ATM – ataxia-telangiectasia mutated kinase signaling pathway

ATR – ATM and Rad50 dependent kinase signaling pathway

BPV1 – bovine papillomavirus type 1

Cdk2 – cyclin-dependent kinase 2

DBD – DNA-binding domain

DDR – DNA damage response pathways

DH complex – E1 double-hexamer complex

dsDNA – double-stranded DNA

DSB – DNA double-strand breaks

DT complex – E1 double-trimer complex

E region – early proteins encoding genomic region

E1BS – E1-binding site in the ori

E1-HA HPV – HPV with HA-tag encoding sequence cloned into E1 ORF

E2F – transcription factors that regulate entry into S-phase

E6-AP – E6-associated protein

EV – rare hereditary disorder *Epidermodysplasia verruciformis*

HA-tag - human influenza hemagglutinin peptide

HD – C-terminal helicase domain of E1

HNC – head and neck cancer

HPV – human papillomavirus

L region – late proteins encoding genomic region

LT – large T-antigen protein

MCS – multiple cloning site

NES – Crm1-dependent nuclear export signal

NGM – normal growth medium

NLS – nuclear localization signal

OBD – origin-binding domain of E1

ori – origin of replication

ORF – open reading frame

p53 – tumor suppressor protein 53

PAE – early polyadenylation site

PAL – late polyadenylation site

PaVE – Papillomavirus Episteme database

PCNA – proliferating cell nuclear antigen

PCR – polymerase chain reaction

Pol  $\alpha$ -prim – DNA polymerase  $\alpha$ -primase complex

pRb – retinoblastoma tumor suppressor protein

PVs – Papillomaviruses

RPA – replication protein A

RT – room temperature

RFC – replication factor C

ssDNA – single-stranded DNA

SSE – stratified squamous epithelium

SUMO – small ubiquitin-like modifier

SV40 – Simian virus 40

TA domain – E2 N-terminal transactivation domain

Topo I – topoisomerase I

URR – noncoding upstream regulatory region

VLPs – virus-like particles

wtHPV – wild-type HPV

# INTRODUCTION

Most of sexually active people have human papillomavirus (HPV) infection during their life (Bucchi et al., 2016). Papillomaviruses (PVs) are widespread in human population, and they are not only pathogenic viruses, but also present in normal skin microbiota (Hazard et al., 2007). They infect epithelial keratinocytes through microinjuries and can persist in skin or mucosa for a long time. The main reason to study HPVs more vigorously is their association with cancerous diseases including anogenital cancers and oropharynx or other head- and neck-associated tumors. For example, the largest part of cervical carcinoma shares the presence of HPV genomic DNA (Adams et al., 2014). Nowadays, vaccination is the main direction to prevent HPV infection (Hancock et al., 2018). On the other hand, HPV-associated cancers cannot be cured with medicines or vaccination, so only surgical operation or other invasive methods can help. Therefore, functional drugs against HPV infections are needed. Better understanding of the HPV life cycle can help in discovering new methods for prevention or treatment of HPV-associated diseases.

HPV usually encodes at least eight functionally different proteins. Among them, the E1 protein is one of the best targets for the development of antiviral drugs. E1 has various functions in the HPV infection cycle, but the main E1 function is participation in HPV DNA replication (Bergvall et al., 2013). E1 and E2 are the only viral proteins that are necessary and sufficient for HPV replication, and inhibiting their activities is a good possibility to inhibit HPV infection at all (Yang et al., 1993). Detection and purification of the E1 protein in laboratory systems is a huge problem, and therefore studies on E1-dependent processes in HPV replication and development of new E1-targeting drugs are also complicated.

The aim of this study is to create a new tool suitable for studies on the E1 protein more sensitively and directionally. The present thesis contains two parts. The first part is the literature overview where HPV types, viral life cycle and models for studying it, HPV protein functions, and E1 protein characteristics are reviewed in details. The experimental part contains the description of the generation of HA-tag-containing HPV types 11 and 18 constructs and the verification of their replication status in the U2OS cell line in comparison with wild-type HPV (wtHPV) types 11 and 18. This study was performed in the molecular virology research group, Institute of Technology, University of Tartu.

# 1. LITERATURE OVERVIEW

## 1.1 Papillomaviridae

Papillomaviruses (PVs) are species-specific, double-stranded DNA (dsDNA) small viruses that infect epithelial keratinocytes. Their icosahedral, non-enveloped virion contains a circular genome approximately 8 kbp in length, which is packed into nucleosomes by cellular histones and usually encodes 8 genes (Favre et al., 1977). PVs infect a wide range of vertebrate species including majority of mammals, but also reptiles and birds (de Villiers et al., 2004). Based on the structure of the skin, the most suitable hosts for PVs are mammals, including humans (Bravo et al., 2010).

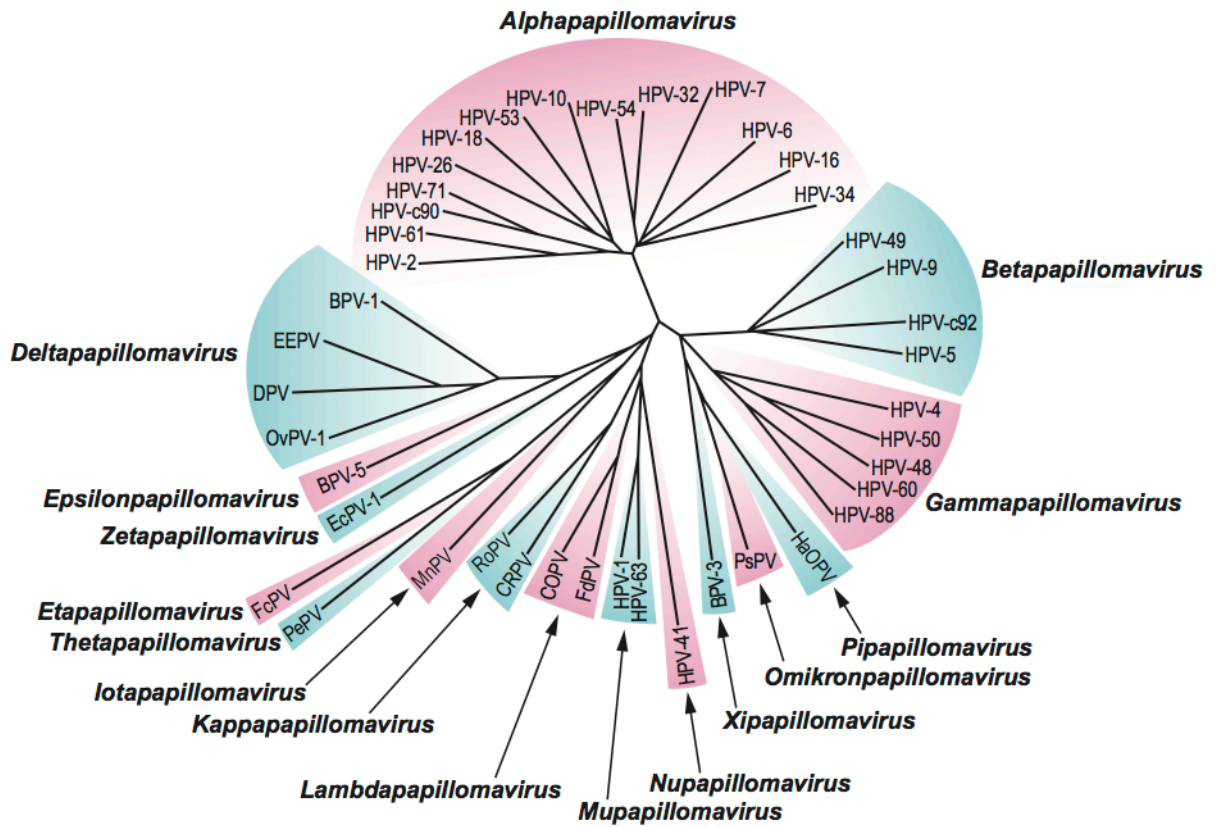
The classification of different types of PVs is based on the nucleotide sequence of the L1 major capsid protein open reading frame (ORF) since it is one of the most conserved sequences in the PV genome. When the virus L1 ORF differs at least 10% from the known types of PVs, the virus is classified as a new type (de Villiers et al., 2004). At present, over 400 PVs belonging to more than 40 different genera of the *Papillomaviridae* family are identified, and due to novel methods of identification, the number of described viruses is continuously increasing (Figure 1). PVs are listed in the fast-growing Papillomavirus Episteme (PaVE) database [1].

## 1.2 Human Papillomaviruses

Human Papillomaviruses (HPVs) are widespread pathogenic DNA viruses. HPVs are phylogenetically distributed into five different families: alpha- ( $\alpha$ ), beta- ( $\beta$ ), gamma- ( $\gamma$ ), mu- ( $\mu$ ), and nu- ( $\nu$ ) (de Villiers et al., 2004) (Figure 1). Currently, more than 200 different types of HPVs are described [1].

HPVs are divided into two groups, depending on their ability to infect cutaneous or mucosal epithelium. Cutaneous and mucosal HPVs are further subdivided into high- and low-risk types according to their potential to induce malignant progression of the lesion in the host organism (Mistry et al., 2008). Low-risk HPVs produce benign hyperkeratotic lesions (warts and condylomas) with a low probability of developing cancer (Egawa & Doorbar, 2017). In contrast, high-risk types are capable of inducing malignant tumors (Bucchi et al., 2016).





**Figure 1.** HPVs belong to alpha- ( $\alpha$ ), beta- ( $\beta$ ), gamma- ( $\gamma$ ), mu- ( $\mu$ ), and nupapillomaviruses ( $\nu$ ). The figure is adapted from the ICTV database [2].

### 1.2.1 Cutaneous HPVs

Cutaneous HPVs belonging to  $\beta$ -,  $\gamma$ -,  $\nu$ -,  $\mu$ -, and some  $\alpha$ -papillomaviruses infect skin (Bzhalava et al., 2013). Low-risk HPVs are well adapted to the host, generally cause asymptomatic infections, and can be found in healthy skin. Therefore, they are considered as a part of the normal microbiota of the human skin (e.g., HPV12 or HPV20) (Hazard et al., 2007). However, some skin HPVs can cause benign epithelial growths, such as warts or papillomas, which disappear after some time due to the adaptive immune system (Amador-Molina et al., 2013).

Nevertheless, high-risk cutaneous HPVs (e.g., HPV3 or HPV10) can induce malignant tumors in immunosuppressed individuals (Reusser et al., 2015). For instance, in the case of an extremely rare autosomal recessive hereditary disorder *Epidermodysplasia verruciformis* (EV), also known as treeman syndrome, the immune system is not able to eliminate HPV infection (e.g., HPV5), and primary lesions like papules and macules form on the skin and can progress to cutaneous carcinomas over time (Leiding & Holland, 2012).

### 1.2.2 Mucosal HPVs

Mucosal HPVs belong to the genus  $\alpha$ -papillomaviruses (Bzhalava et al., 2013). Some mucosal HPVs belong to high-risk types because of their oncogenic potential (HPV types 16, 18 or 33), and therefore, they are the most studied PVs. Mucosal HPVs are mostly associated with anogenital area cancers, but also with upper respiratory tract, oral cavity, and other head and neck tumors. For instance, high-risk mucosal HPV16 prevails in HPV-associated head and neck squamous cell carcinomas and also is found in cervical carcinomas (50% of cases). HPV18 is another HPV most commonly identified in cervical carcinomas (20% of cases) (Adams et al., 2014).

Low-risk mucosal HPVs (e.g. HPV6 and HPV11) produce benign tumorous warts, called papillomas (condylomas in the anogenital area), which in most cases resolves spontaneously, and their progression to cancer is not generally observed in healthy people (Chow & Broker, 2013). Some HPVs are classified as probable high-risk types (e.g., HPV types 26, 53 and 66) (Bucchi et al., 2016). High-risk HPV18 and low-risk HPV11 studied in the present work belong to mucosal  $\alpha$ -papillomaviruses.

### 1.2.3 Prevention or treatment of HPV infection

HPVs as the most common sexually transmitted viral pathogens cause approximately 5% of the global cancer burden (Egawa & Doorbar, 2017). HPVs are mostly associated with cervical cancer and other anogenital cancers including anal, vulvar, and penile cancers (Chabeda et al., 2018). Although the main route of HPV infection is sexual contact, nonsexual transmission can occur from mother to child prenatally or perinatally (Lee et al., 2013). Recently, high-risk HPV infections were found to be also associated with head and neck cancers (HNCs), especially in the oropharynx (Ang et al., 2011). Approximately 98% of cervical carcinomas and at least 25% of HNCs share the presence of HPV DNA (Adams et al., 2014). In developed countries, cervical cancer screening allows early detection of diseases resulting in overall decrease of cervical cancer mortality. Unfortunately, cervical cancer is much more common in developing countries. Therefore, it is the second leading cause of cancer-related death in women in the developing world (Ladner et al., 2016).

Now, there are three HPV vaccines approved to use. The Gardasil induces antibody production against high-risk HPV16 and HPV18 and protects against low-risk HPV6 and HPV11. Cervarix vaccines are targeted only against HPV16 and HPV18. Recently, Gardasil 9 or nine-valent vaccine against HPV types 6, 11, 16, 18, 31, 33, 45, 52, 58 has been introduced to the market. Available vaccines are successful, but there are lots of additional high-risk HPV types

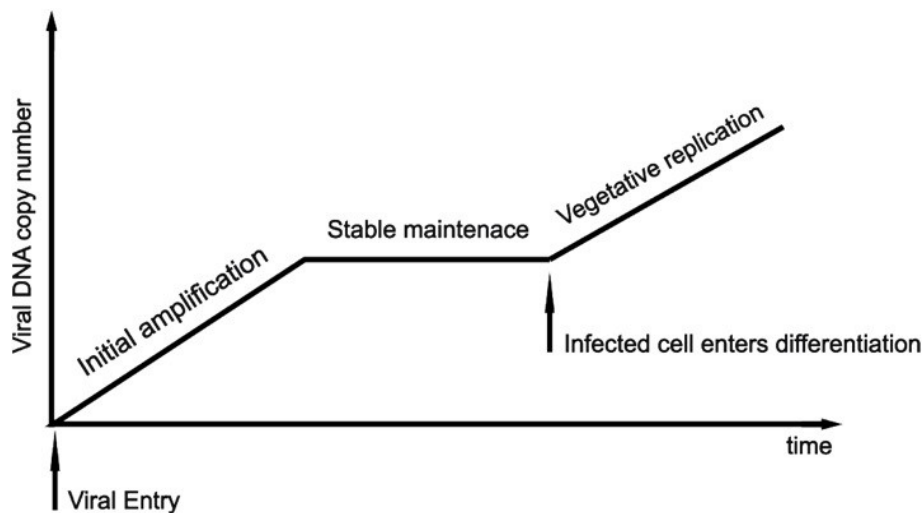
associated with about 30% of cervical cancers (Hancock et al., 2018). Besides, vaccination does not cure people already harboring HPV infection. Therefore, studies of HPV molecular biology and life cycle are essential for HPV-associated cancer prevention and treatment (Winters et al., 2006).

### 1.3 HPV infection cycle

HPVs infect stratified squamous epithelium (SSE) (Kajitani et al., 2012). This epithelium is subdivided into two types: keratinized SSE (e.g., epidermis of the skin) and nonkeratinized SSE (e.g., mucosa of vagina or anal canal). Accordingly, HPVs are also divided into cutaneous and mucosal types.

The main part of all SSEs is keratinocytes at various stages of differentiation. HPVs infect highly mitotic basal keratinocyte stem cells located in the basal layer of SSE because of their ability to accomplish the cycle of dividing and differentiation (Longworth & Laimins, 2004). These cells divide asymmetrically: one daughter cell remains in the basal layer as a stem cell of the epithelium; the other moves to the upper layers undergoing the programmed differentiation. The HPV infection cycle is roughly connected to differentiation of keratinocytes.

HPV enters the host cell in the basal layer through the epithelial microinjuries. During the life cycle, HPV undergoes three different replication phases that reflect the stages of the infection (Figure 2) (McBride, 2008). During the first stage named transient initial amplification, HPV copy number reaches about 50–100 per cell (Kadaja et al., 2009). The second stage is stable maintenance replication during latent infection, and the HPV copy number remains stable. These phases occur in the epithelial basal layer (McBride, 2008). The last step is vegetative amplification, which occurs during the productive phase of the infection, when the HPV amount grows to more than 1000 copies per cell. This process takes place in terminally differentiated keratinocytes when virions are starting to form (Kadaja et al., 2009).

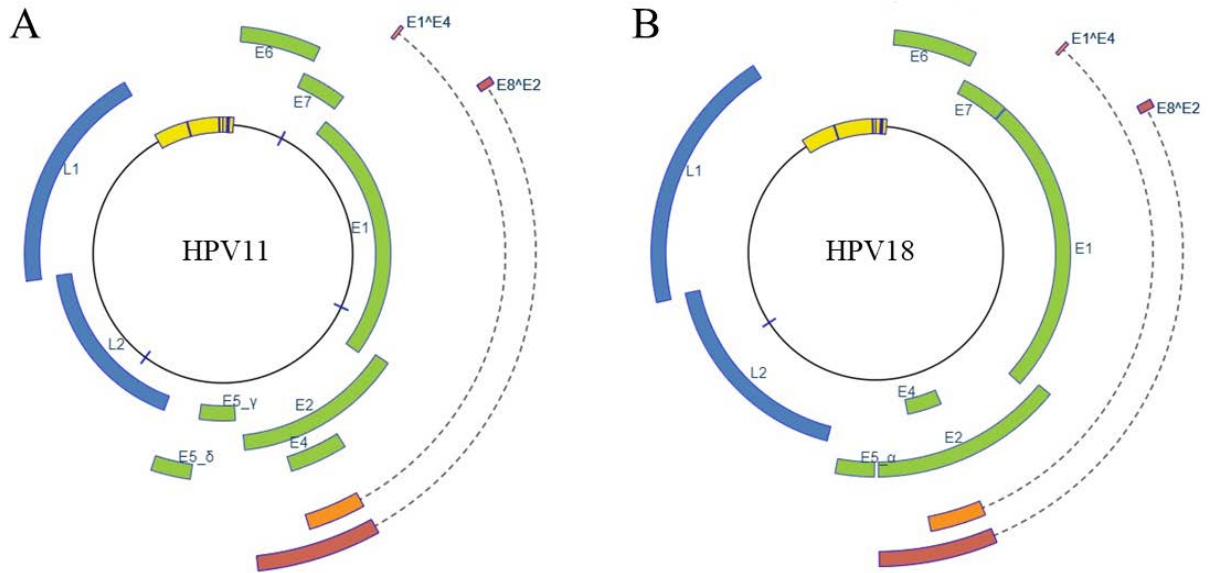


**Figure 2.** Representation of the HPV life cycle. HPV enters the cell and amplifies itself in the basal layer to the stable state. During programmed differentiation of keratinocytes, HPV amplifies to around 1000 copies per cell and eventually forms virions that leave the terminally differentiated cell (Ustav jr., 2016.).

## 1.4 HPV molecular biology

### 1.4.1 HPV genomic organization

The genome organization of different HPVs is generally similar. A schematic representation of HPV11 and HPV18 genomes is shown in Figure 3. The HPV genome is an approximately 8 kbp dsDNA circular molecule with open reading frames (ORFs) located on a single DNA strand (Graham, 2010). The genome can be divided into two coding regions: an early (E) and a late (L) regions and a noncoding upstream regulatory region (URR), located between them. The transcripts from the E region are produced during almost all stages of the viral life cycle until packaging of the genome in the upper layers of epithelium. Expression of late ORFs encoding L1 and L2 capsid proteins occurs only at the end of the virus infection cycle when the viral genome has to be packed (Graham, 2010). The noncoding URR region contains binding sites for the main viral regulatory protein (E2) and cellular transcription factors, the origin of the replication (ori), and the early and late promoters (McBride, 2008). In some HPVs (e.g., HPV11 or HPV16), the late promoter is located in the E7 ORF (Grassmann et al., 1996). There are also early (PAE) and late (PAL) polyadenylation sites, in which the synthesis of early or late transcripts, respectively, ends (Graham, 2010). As PV mRNAs are polycistronic, it is difficult to identify all transcripts that might have a role in the viral life cycle (Wang et al., 2011).



**Figure 3.** Schematic representation of HPV11 and HPV18 genomes. HPV genome can be divided into three different regions: early ORFs (E1, E2, E4, E5, E6, and E7) indicated by green, late ORFs (L1 and L2) shown in blue, and URR region (yellow). Also HPV encodes some truncated forms of proteins using different ORFs (E1<sup>E4</sup> and E8<sup>E2C</sup>) (orange and pink). The URR contains different regulatory sequences and binding sites for viral and cellular proteins. The figure is adapted from the PaVE database [1].

### 1.4.2 Overview of viral proteins

HPV pre-mRNAs transcribed from the early region are alternatively spliced producing a variety of functional mRNAs that encode proteins E1, E1<sup>E4</sup>, E2, E8<sup>E2C</sup>, E5, E6, and E7. These proteins are required for HPV replication, transcription, genome segregation, and modulation of the host cell cycle (Graham, 2010). The L1 and L2 proteins, synthesized from the late coding region, form the virus capsid and have a role in HPV interactions with a host cell (Aksoy et al., 2017).

#### 1.4.2.1 Early region proteins

**The E1 protein** (approximately 68 kDa) is absolutely needed for replication of the viral genome (Blitz & Laimins, 1991). It is the only enzyme encoded by HPVs (Yang et al., 1993). The E1 proteins of HPV11 and HPV18 are in the focus of this thesis. Therefore, the structure of E1 and its role in the HPV life cycle are discussed in more detail in chapters 1.5 and 1.6 of this thesis.

Fusion **protein E1<sup>E4</sup>** is considered to be an early protein, but during the HPV life cycle, E1<sup>E4</sup> generally becomes detectable at the beginning of the late stages of the infection cycle (Wilson et al., 2007). In the early stages of the infection, the synthesis of the E1<sup>E4</sup> protein occurs at a very low level. Nevertheless, the E1<sup>E4</sup> protein is believed to be implicitly involved

in replication and transcription of viral DNA via binding and stabilizing the E2 protein (Davy et al., 2009). Also, it has been suggested that the E1<sup>E4</sup> protein serves to optimize the exit of virions from cells in the upper epithelial layers (Doorbar, 2013). In addition, the E1<sup>E4</sup> protein is capable to induce a cell cycle arrest in G<sub>2</sub> phase and to target the host cell to apoptosis (Davy et al., 2002).

**The full-length E2 protein** is a transcription factor consisting of two domains linked together by a linker sequence called the hinge region. The E2 C-terminal region is a DNA-binding domain (DBD). On the other hand, the N-terminal region forms a transactivation (TA) domain (McBride, 2013). E2 is a multifunctional protein that is involved in various HPV life cycle processes. E2 interacts with different cellular proteins and plays a key role in controlling the transcription of viral proteins (Graham, 2016). Transcriptional activation or repression depends on the binding affinity of the full-length E2 protein for its binding sites (ACCGN<sub>4</sub>CGGT motifs) located in the URR region (Vösa et al., 2012). At lower concentrations, the E2 protein acts as a transcriptional activator, higher concentrations of E2 lead to repression of the viral DNA transcription (Bouvard et al., 1994). The E2 protein also plays an important role in initiation of viral DNA replication and segregation of HPV genomes in mitotic cells (McBride, 2013). The E2 protein of high-risk HPVs is capable of activating apoptosis in host cells, which can be important in a context of releasing viral particles in environment (Singh et al., 2016). Also, some studies indicated that the E2 protein enhances packaging of HPV genomes in virions through interactions with the L2 protein (Zhao et al., 2000). Disruption of E2 during the process of HPV DNA integration into the host cell genome may also lead to increased oncogenic potential of the viral E5, E6, and E7 proteins (Williams et al., 2011).

All PVs have been shown to express an **E8<sup>E2C</sup>** (or equivalent) **inhibitor protein** that is encoded by a truncated E2 transcript (Straub et al., 2014). The E8<sup>E2C</sup> protein lacks a TA domain but possesses the E8 ORF-encoded sequence, which is located in the E1 ORF, a C-terminal DBD and a linker-region of E2 (McBride, 2013). E8<sup>E2C</sup> is the only protein that can compete with full-length E2 for binding sites in the URR region and repress early transcription and replication. Knock-down of E8 ORF results in increased copy numbers of the HPV genome. Also E8<sup>E2C</sup> can form an inactive heterodimeric form with the E2 protein repressing transcription and replication (Straub et al., 2014).

Generality of PVs encodes the **E5 protein**. High-risk HPV types, including HPV18, encode one E5 $\alpha$  protein. Low-risk HPV types such as HPV11 encode two E5 proteins: E5 $\gamma$  and E5 $\delta$  (Figure 3) (DiMaio & Petti, 2013). High-risk and low-risk HPV E5 proteins are different. Nevertheless, their expression correlates with higher risk of cancer suggesting that E5 proteins

can enhance the transforming properties of E6 and E7 oncogenic proteins (Bravo & Alonso, 2004). Also, the E5 protein interacts with different cellular factors, but their role in the HPV life cycle is poorly understood (DiMaio & Petti, 2013).

The main function of the **E6 protein** is to restore the proliferation of differentiated cells (Oh et al., 2004). When a cell enters the unplanned S-phase, it is typically directed to apoptosis. The HPV E6 oncogenic protein binds to a cellular E6-associated protein (E6-AP), and the complex E6/E6-AP binds to major tumor suppressor protein 53 (p53), which results in a rapid ubiquitin-dependent degradation of p53 (Yim & Park, 2005). E6 shows the ability to block other apoptotic pathways as well, so abnormally active mitotic cells can support viral replication (Howie et al., 2009).

The **E7 protein** acts in cooperation with the E6 protein, but targets retinoblastoma (pRb) tumor suppressor protein (Moody & Laimins, 2010). pRb protein in complex with the E2F transcription factor regulates the entry into the S-phase of the cell cycle, when a cell genome replicates. When the E7 protein binds to the pRb protein, pRb is directed to degradation and the transcription factor E2F is released from the inhibitory complex (Yim & Park, 2005). Next, the host cell division and DNA replication are activated enabling simultaneous replication of the HPV genome. The oncogenic potential of E5, E6, and E7 of high-risk HPVs is higher due to their higher affinity for cellular tumor suppressor proteins (Moody & Laimins, 2010).

#### 1.4.2.2 Late region proteins

The icosahedral, non-enveloped HPV virions consist of two types of proteins: **major capsid protein L1** and **minor capsid protein L2**, whereas only the L1 protein is present on the HPV surface (Buck et al., 2013). The synthesis of capsid proteins occurs at the final stages of the virus life cycle followed with packaging and release of virions into the environment. It is assumed that L1 participates in primary interaction of HPV with the host cell and assists viral genome trafficking in the nucleus (Letian & Tianyu, 2010). L2 is believed to play an important role in the encapsidation of the circular and nucleosome-bound viral genome during assembly of the virions in the nucleus of the host cell (Wang & Roden, 2013). Recombinant L1 protein self-assembled virus-like particles (VLPs) is the main direction of the development of vaccines against HPV infections (Wang & Roden, 2014). In contrast to L1, the L2 protein cannot assemble into VLPs, but L2 co-assembled with L1 enhances stability of the produced VLPs (Wang & Roden, 2013).



## 1.5 E1 protein structure

The E1 protein is encoded by the largest and one of the most conserved ORFs in the PV genome, which is located near the URR region after the E6 and E7 ORFs (Figure 3) (Bergvall et al., 2013). A lot of understanding of the E1 protein structure and functions came from a similar, well-studied large T-antigen (LT) protein that belongs to Simian virus 40 (SV40) from the family *Polyomaviridae* (Clertant & Seif, 1984). *Polyomaviridae* and *Papillomaviridae* viruses use host cell DNA replication machinery in a similar manner for replication of their own genomes.

The E1 protein is about 600-660 bp in length (Bergvall et al., 2013). It can be divided into three functional domains. The N-terminal part is the least conserved domain of the E1 protein (Amin et al., 2000). The central DBD binds to the specific sequence of the ori (Titolo et al., 2003). The C-terminal region possessing 3'-5' helicase activity has many functions in the active enzymatic form, and is responsible for unwinding short DNA duplexes (Sedman & Stenlund, 1998). A schematic representation of domains is shown in Figure 4.

**The N-terminal domain** is about 200 amino acids. There are different motifs that are changeable in their sequence, but present practically in all PV types (Bergvall et al., 2013). These are bipartite nuclear localization signal (NLS), a Crm1-dependent nuclear export signal (NES), a cyclin-binding motif that interacts with cyclin A/E in complex with cyclin-dependent kinase 2 (Cdk2), and phosphorylation sites for Cdk2 and other kinases (Bergvall et al., 2013). All of them are essential for correct replication *in vivo*, but unnecessary for replication *in vitro* (Amin et al., 2000).

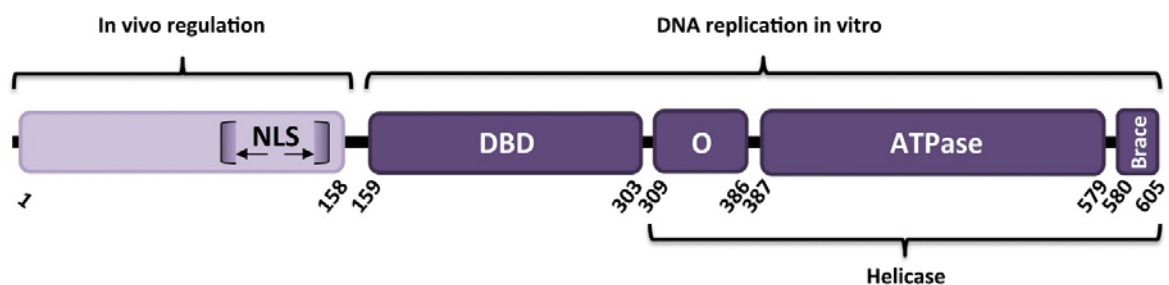
**The origin-binding domain (OBD) or DBD** is needed to recognize specific sequences in the ori of viral DNA (Thorner et al., 1993). PVs have six E1-binding sites (E1BS1 to E1BS6) of consensus sequence 5'-ATTGTT-3' in the E1 binding region (Titolo et al., 2003). Some of these sites can be degenerated or suffer from lower affinity for E1 (Chen & Stenlund, 2001). The DBD region in general has low affinity for E1BSs, and the E1-E2 complex is required for effective binding of E1 to the ori (Thorner et al., 1993). Even dimerization of E1 proteins increases the affinity of the DBD for its consensus sequence by around 10 times (Titolo et al., 2003).

The enzymatic **C-terminal helicase domain (HD)** is required for correct unwinding of DNA during the replication. Also, it interacts with the TA domain of the E2 protein when the pre-helical complex is forming at the ori (Yang et al., 1993). The HD can be also subdivided into three different functional regions (Figure 4). These are the minimal oligomerization domain,



which participates in the formation of the correct hexameric complex, the AAA+ superfamily domain with ATPase activity, and the C-terminal brace (Sedman & Stenlund, 1998). The role of the C-terminal brace is not fully understood, but it may be important for the stabilization of the oligomeric state of E1 when the enzymatic E1 helicase complex faces conformational changes during the binding and hydrolysis of ATP (Whelan et al., 2012).

The helicase activity of HPV is absolutely needed for correct replication of the HPV genome. Therefore, it is a good target for antiviral drugs. Only a few drugs inhibiting ATPase activity of the hexameric E1 complex or repressing the interactions between E1 and E2 have been found. This inhibition results in lower levels of HPV replication yielding a lower amount of HPV genomes per cell (Archambault & Melendy, 2015).



**Figure 4.** Schematic representation of E1 protein structure. E1 protein can be subdivided into three regions. The N-terminal domain is essential for viral DNA replication and contains regulatory sequences. The DBD is located in the center. It has low binding specificity to viral genome and requires the E2 protein to initiate replication. The C-terminal domain has three parts: oligomerization domain (O), which has an important role in forming active helicase complex, ATPase domain, and C-terminal brace responsible for stabilization (Bergvall et al., 2013).

## 1.6 Role of E1 in the life cycle of HPVs

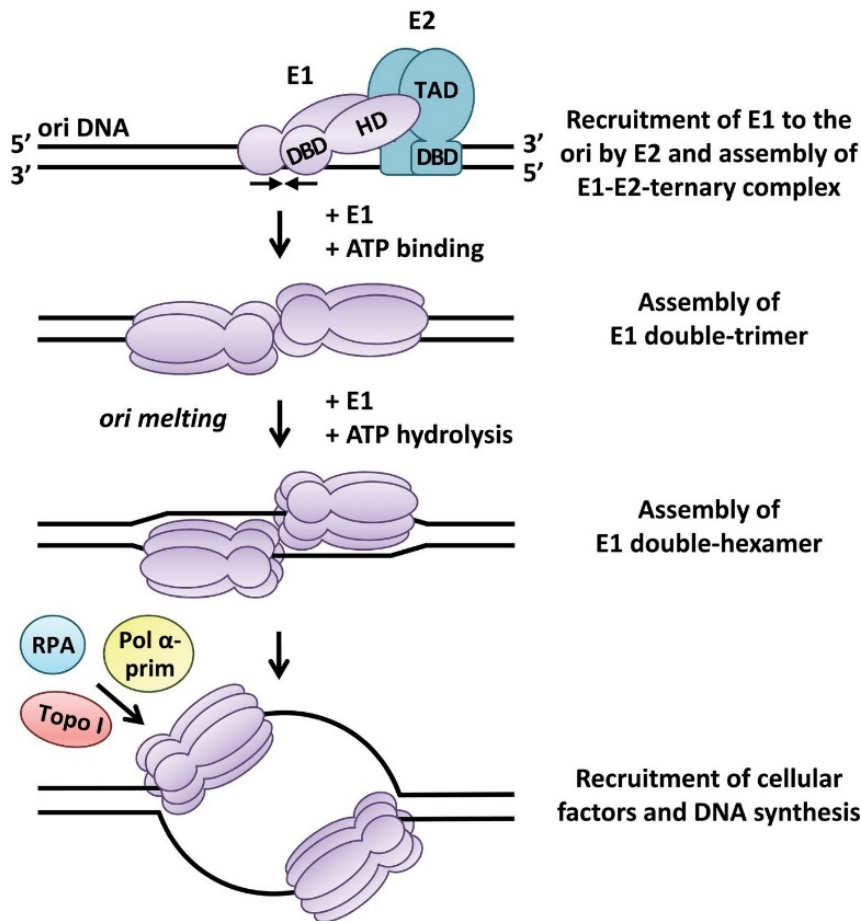
The main function of the E1 protein during the HPV life cycle is to initiate the replication of the HPV genome, which is further accomplished mainly with the help of the host cell replication machinery. E1 and E2 are the only viral proteins required for HPV replication (Yang et al., 1993). Also, the E1 protein interacts with several cellular factors needed for replication of the HPV genome, like DNA polymerase  $\alpha$ -primase (Pol  $\alpha$ -prim) complex, replication protein A (RPA) or topoisomerase I (Topo I) (Melendy et al., 1995). Moreover, the E1 protein can produce double-strand DNA breaks (DSBs) in the host cell genome, which results in the

activation of DNA damage response (DDR) pathways and induces recombination-dependent replication of HPV DNA (Sakakibara et al., 2011).

### 1.6.1 E1 and replication of the HPV genome

Since E1 itself can poorly bind to DNA, it cooperates with the E2 protein that binds with high affinity to four E2-specific palindromic motifs located at the ori (Ustav et al., 1993). The E2 protein assists loading of E1 specifically at the ori and prevents nonspecific interactions between the DBD of E1 and cellular DNA. Their interaction occurs via the E2 N-terminal TA domain and the E1 C-terminal region (Berg & Stenlund, 1997).

ATP is allosteric effector of E1-E2 complex dissociation. ATP promotes the oligomerization of the E1 hexameric complex. ATP and the E2 TA domain compete with each other for binding to the E1 C-terminal HD, but they interact with different motifs of the E1 helicase region. When ATP is present near the ori, it can bind to the E1-E2 complex and promote its transition to the E1 double-trimer (DT) complex (Bergvall et al., 2013). The E1 DT complex transforms into the double-hexamer (DH) complex, which forms a ring-like structure and is known as the active enzymatic form of the E1 protein that can perform bi-directional melting of the viral genome at the ori. Each E1 active enzymatic hexamer complex encircles one DNA strand (Sedman & Stenlund, 1998). A schematic representation of this process is present in Figure 5.



**Figure 5.** Assembly of the active enzymatic E1 complex. Firstly, E1 dimer binds to E1BS at the ori with the help of E2 proteins. ATP sets free E2 protein molecules, and the E1 double-trimer (DT) complex is formed. Thereafter, a subsequent helicase double-hexamer (DH) complex forms, and replication can be started with the help of cellular factors (Bergvall et al., 2013).

### 1.6.2 Interactions between E1 and host cellular factors

E1 plays an important role not only in replication of the HPV genome, but also in a wide spectrum of other viral life cycle processes. E1 is responsible for interactions with a huge variety of cellular proteins, some of them are p80/Uaf1, Cyclin A/E-Cdk2, Ini1/hSNF5, Hsp40, Hsp70, Histone H1, E1-BP/TRIP13, and Importins (Bergvall et al., 2013).

E1 interacts using the DBD and the HD with Topo I, which is an ATP-independent protein that decreases torsional stress in DNA and modifies the topology of supercoiled DNA, which results in increased binding of E1 to the ori and faster assembly into the active enzymatic form (Clower et al., 2006). The next important protein is replication protein A (RPA), which also interacts with the E1 DBD. During the replication, E1-bound RPA molecules dissociate from E1 and bind to naked single-stranded DNA (ssDNA) of the lagging strand (Loo & Melendy, 2004).

Also, E1 interacts with the DNA polymerase  $\alpha$ -primase (Pol  $\alpha$ -prim) complex (Park et al., 1994). The Pol  $\alpha$ -prim complex makes RNA-DNA primers, which are needed for DNA Polymerase  $\delta$  to synthesize a total DNA sequence of the lagging and leading strands (Prindle & Loeb, 2012). Interestingly, Pol  $\alpha$ -prim binds to E1 only after the dissociation of the E1-E2 complex (Amin et al., 2000). To prevent re-association of Pol  $\alpha$ -prim *in vitro*, HPVs need replication factor C (RFC) in cooperation with RPA molecules. RFC also loads the proliferating cell nuclear antigen (PCNA) and DNA polymerase  $\delta$  for the synthesis of remain DNA (Melendy et al., 1995). Interactions of the E1 protein or other viral proteins with these cellular factors have not been reported *in vivo* (Bergvall et al., 2013).

### 1.6.3 E1 toxicity for the host cell and regulation of E1

E1 is a very toxic protein for host cell DNA. Therefore, the levels of E1 are regulated tightly by synthesizing a very low amount of E1 or by its degradation (Sakakibara et al., 2011). Studies with bovine papillomavirus type 1 (BPV1) E1 showed that E1 forms the complex with cyclin E/Cdk2 that stabilizes E1 before HPV genomic replication (Cueille et al., 1998). However, E1 becomes destabilized again during the replication cycle and is degraded with the help of the ubiquitin-proteasome pathway, which needs poly-ubiquitination of the E1 protein by the anaphase-promoting complex/cyclosome (Mechali et al., 2004).

E1 is regulated by different post-translational modifications. For instance, subcellular localization of E1 is regulated by phosphorylation (Lentz et al., 1993). E1 is also subjected to sumoylation or covalent attachment of a small ubiquitin-like modifier (SUMO) that is suggested to affect the enzymatic activity of E1 (Fradet-Turcotte et al., 2009). During the vegetative replication, HPV activates caspases-3, -7, and -9 that promote specific cleavage of E1 (Bergvall et al., 2013). However, it is currently unknown how E1 proteolytic cleavage helps with the genome amplification during the productive phase (Richard & Tulasne, 2012).

### 1.6.4 E1-induced DNA damage response

Nuclear accumulation of the E1 active enzymatic form arrests the cell cycle in the S-phase and blocks the proliferation of the host cell (Fradet-Turcotte et al., 2011). The E1 NLS in combination with the NES forms a pathway that delicately controls the levels of replication of the HPV genome during the S-phase in the host cell nucleus (Fradet-Turcotte et al., 2010).

E1 active helicase always tries to initiate host DNA replication thereby inducing cellular DNA DSBs that activate the DDR pathways in the host cell. E2 in complex with E1 attenuates the ability of inducing DSBs, but does not fully prevent it, so the DDR takes place during HPV genome replication (Fradet-Turcotte et al., 2011). The main DDR pathways are ataxia-

telangiectasia mutated (ATM) kinase signaling pathway and ATR (ATM and Rad3-related) pathway (Reinson et al., 2013). These pathways can activate homologous recombination, and virus can use it to replicate its own genome during the vegetative amplification stage (Moody & Laimins, 2009).

E1-dependent DBSs can induce host genomic instability and promote the integration of the HPV genome into the host cell genome (Moody & Laimins, 2010). Integration of high-risk HPV genomes is one of the factors that can lead to cancer (Pett & Coleman, 2007).

## 1.7 Difficulties to study the HPV life cycle

The HPV infection cycle is closely connected with programmed differentiation of keratinocyte stem cells. Because of this fact, most of eukaryotic cell lines do not support HPV replication and are unsuitable for studying the HPV life cycle. One method to study the HPV life cycle is organotypic raft cultures. They are either derived from biopsies (primary human keratinocytes or cervical carcinoma dysplasia) or created by transfection of keratinocytes with HPV DNA (Anacker & Moody, 2012). High cost and low efficiency are the raft culture disadvantages since low-risk HPVs cannot immortalize keratinocytes and their transcripts are poorly detected.

There are some HPV-positive cell lines available. For instance, the HPV31b+ CIN 612 cell line is derived from human cervical intraepithelial neoplasia (CIN), HPV16+ W12 cell line is established from epithelial cells of low-grade cervical lesion, and HPV16+ KG cell line is designed from a vulvar intraepithelial neoplasm (De Geest et al., 1993; Doorbar et al., 1990; Grassmann et al., 1996). All these cell lines are suitable for studying the life cycle of their specific HPV types only.

The molecular virology research group (Institute of Technology, University of Tartu) has developed a cellular system for studying different HPV types (Geimanen et al., 2011). It is based on the U2OS cell line derived from human osteosarcoma. The greatest benefit of this system is that it can support replication of not only high-risk HPVs like HPV16 and HPV18, but also low-risk HPVs (HPV6b and HPV11) and cutaneous types (HPV5 and HPV8). Although the U2OS cell line does not allow studying the entire HPV life cycle from entry to the cell to encapsidation of viral genomes, it perfectly supports the study of all three HPV replication phases. The U2OS cell line is used also in the present study.

## 2. EXPERIMENTAL PART

### 2.1 Aims of the study

The general purpose of this study is to create new E1-HA HPV11 and E1-HA HPV18 constructs by cloning the HA-tag encoding sequence in E1 ORF of the viral genomes with the subsequent goal to use the constructs in further studies. Step-by-step aims of the work are:

- Synthesis of viral DNA fragments containing HA-tag within them;
- Cloning of the synthesized DNA fragments into the HPV parental genomes;
- Comparison of the replication status of wtHPVs and the created E1-HA HPV constructs in the U2OS cell line.

## 2.2 Materials and methods

### 2.2.1 Plasmids used in the present study

- ❖ **pJET1.2/blunt Vector** (2974 bp). This is a vector designed for cloning blunt-ended polymerase chain reaction (PCR) products (CloneJET PCR Cloning Kit, *Thermo Fisher Scientific*). It contains the  $\beta$ -lactamase *bla* gene, which gives resistance to ampicillin ( $\text{Ap}^R$ ), specific primer binding sites, and a multiple cloning site (MCS).
- ❖ **pMC.BESPX** (4084 bp). This vector contains the following elements necessary and sufficient for production of minicircle plasmids: attB and attP sites, where bacteriophage  $\Phi\text{C31}$  integrase can bind, 32 restriction sites of SclI endonuclease, MCS, kanamycin resistance gene ( $\text{Kan}^R$ ), and the origin of bacterial replication (ColE1 ori) (Kay et al., 2010).
- ❖ **HPV11-pMC.BESPX** (12015 bp). This plasmid was constructed to produce HPV11 minicircle plasmids by cloning wtHPV11 genomic DNA into the BamHI site of the pMC.BESPX vector (Orav et al., 2015).
- ❖ **HPV11 minicircle DNA plasmid** (7967 bp). Covalently closed, supercoiled, circular plasmid produced from **HPV11-pMC.BESPX**. It differs from wtHPV11 DNA by only 36 bp in length attR sequence in its URR that was formed during the process of recombination by integrase  $\Phi\text{C31}$ .
- ❖ **HPV18-pMC.BESPX** (11974 bp). This plasmid contains wtHPV18 genomic DNA cloned into the pMC.BESPX vector between attP and attB sites (Orav et al., 2013).
- ❖ **HPV18 minicircle DNA plasmid** (7893 bp). This plasmid was produced analogously to the previously described HPV11 minicircle DNA plasmid.

### 2.2.2 Generation of E1-HA HPV11/HPV18-pMC.BESPX constructs

#### 2.2.2.1 PCR

To synthesize and amplify HPV fragments containing the HA-tag encoding sequence, I used PCR and primers listed in Table 1. The synthesis was performed using Pfu DNA polymerase (*Thermo Fisher Scientific*), HPV11/HPV18-pMC.BESPX DNA constructs as templates for amplification and PCR Mastercycler (*Eppendorf Scientific*). The PCR components are listed in Table 2. The first step of the PCR program was initial denaturation of template DNA for 3 min ( $95^\circ\text{C}$ ). The second part of the reaction comprised 20 cycles of the synthesis: 30 s of denaturation ( $95^\circ\text{C}$ ), 30 s of annealing of the primers ( $57^\circ\text{C}$ ), and 4 min of extension ( $72^\circ\text{C}$ ). The last step was the final extension for 5 min ( $72^\circ\text{C}$ ) and cooling to  $10^\circ\text{C}$ .

**Table 1.** Primers used during the work. The underlined sequence encodes HA-tag. The primers contain the following restriction sites depicted as indicated in the parentheses: NheI (green), BshTI (red), SwaI (yellow), Bpu1102I (grey), and AatII (pink). Sequences complementary to viral genomes are shown in italic.

#	Primer:	Sequence:
1	HPV11 E1-HA BshTI FW	<b>ACCGGT</b> <i>TTCGGTTACCCACACCC</i>
2	HPV11 E1-HA NheI REV	<b>GCTAGC</b> <i>GTAATCTGGAACATCGTATGGGTATG</i> <i>AATCGTCCGCCATC</i>
3	HPV11 E1-HA NheI FW	<b>GCTAGC</b> <i>GGTACAGAAAATGAGGG</i>
4	HPV11 E1-HA SwaI RV	<b>ATTTAAAT</b> <i>TGACTGTCAGCCAAAC</i>
5	HPV18E1-HA Bpu1102I FW	<b>GCTCAGC</b> <i>CAGACGACCTTCGAGC</i>
6	HPV18E1-HA NheI RV	<b>GCTAGC</b> <i>GTAATCTGGAACATCGTATGGGTATT</i> <i>CTGGATCAGCCATTG</i>
7	HPV18E1-HA NheI FW	<b>GCTAGC</b> <i>GGTACAGACGGGGAGG</i>
8	HPV18E1-HA AatII RV	<b>GACGTC</b> <i>TGGCCGTAGGTCTTTGC</i>
9	Sequencing HPV18 E1 seq	<i>CCTGTTGCGTTACAAATGTGG</i>

**Table 2.** Reagents for the PCR.

Reaction components:	Volume (μl):
10x Pfu Buffer with MoSO <sub>4</sub>	5
dNTP mix, 2 mM each	5
Forward primer 10 μM	5
Reverse primer 10 μM	5
Template DNA, 18 ng/μl	2
Pfu DNA polymerase	1
MQ water	27
Total volume	50

### 2.2.2.2 Gel electrophoresis

I analyzed the obtained PCR products using the method of gel electrophoresis. After the PCR, I loaded the mixes on 2% agarose gel containing 0.3 μg/ml of ethidium bromide in the presence of the following DNA ladders: 0.6 μg of O'GeneRuler 1kb and 0.8 μg of pUC19 DNA MspI



(HpaII) (both *Thermo Fisher Scientific*). The gel ran for 45 min in 1xTAE (40 mM Tris-acetate, 1 mM EDTA) buffer under 100 V. After separation of fragments on the gel, I cut out the required fragments with scalpel under UV (260 nm) and put them into clean 1.5 ml tubes.

### 2.2.2.3 DNA extraction from the gel

Each fragment was carefully extracted from the gel using QIAEX II Gel Extraction Kit (*Qiagen*) and according to manufacturer`s instructions. Gel pieces with DNA were dissolved in a QIAEX II solution at 50 °C and loaded to the columns. After centrifugation with Biofuge pico (*Heraeus*) at room temperature (RT) and 5000 rpm, DNA remained bound to silica resin and the flow-through was removed. The column-bound DNA was washed with QX1 and PE buffers and then eluted in 10 µl of pure water. The concentrations were measured using a Nanodrop-1000 spectrophotometer (*Thermo Fisher Scientific*) at 260 nm wavelength.

### 2.2.2.4 Cloning of the PCR products into the vector

The purified PCR products were cloned into the pJET1.2/blunt cloning vector. The ligation reaction components are listed in Table 3. The reactions were performed at RT for 1 h.

**Table 3.** Ligation reaction of the purified PCR products and cloning vector.

Reaction components:	Volume (µl):
10x reaction Buffer	1
pJET1.2/blunt cloning vector (50 ng/µl)	0.3
T4 DNA Ligase	1
Purified PCR product	7.7
Total volume	10

### 2.2.2.5 Transformation in DH5α

For the transformation, I used competent cells of nonpathogenic *Escherichia coli* strain DH5α, which were stored at −70 °C. Prior the transformation, competent cells were thawed on ice for 20 min. Each of four separate 10 µl ligation mixes were added to 200 µl of competent cells and incubated on ice for 30 min. After that, bacteria were incubated at 37 °C for 3 min and then transferred on ice for 1 min. Then I added 850 µl of LB broth (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl) to each tube and incubated them at 37 °C for 45 min. The tubes were centrifuged for 5 min at 5000 rpm and RT, and then supernatant was removed. Bacteria were resuspended in 100 µl of LB broth and plated using the spread plate technique on LB agar supplemented with 50 µg/ml of ampicillin. The plates were incubated at 37 °C for 18 h. After that, several

colonies were transferred from plates to 3 ml of LB broth containing 200 µg/ml of ampicillin and incubated at 37 °C, 220 rpm for 18 h.

#### 2.2.2.6 Plasmid DNA extraction

Plasmid DNA was extracted from bacteria using Plasmid Extraction Mini Kit (*Flavoprep*) according to manufacturer's protocol. All procedures were performed at RT. First, I centrifuged bacteria at 5000 rpm for 5 min. After removal of supernatant, the pellet was resuspended in FAPD1 buffer containing RNase A. Bacteria were lysed in FAPD2 buffer, and then reactions were neutralized with FAPD3 buffer. After centrifugation at 13 000 rpm for 5 min, supernatant containing DNA was transferred to the FAPD column and centrifuged. The column-bound DNA was washed with W1 and Wash buffers, and then eluted with 40 µl of nuclease-free water.

#### 2.2.2.7 Precipitation of DNA

If the concentrations of the obtained DNA were low, I precipitated DNA using 1/10 V (volume of the DNA sample) of 5 M NaCl and 2 V of 96% ethanol per 1 V of sample. The samples were mixed by vortexing and centrifuged at 4 °C and 15000 rpm for 20 min using MicroCL 21R Microcentrifuge (*Thermo Fisher Scientific*). After removal of supernatant, the pellets were washed with cold (−20 °C) 75% ethanol and centrifuged for 7 min at 4 °C and 15000 rpm. Then, the pellets were dried and resuspended in 10 µl of nuclease-free water.

#### 2.2.2.8 Control of the cloned PCR products with restriction

The restriction analysis was used to control the presence and size of our PCR products in the pJET1.2/blunt cloning vector. I used restriction enzymes specific for the sites encoded by 5'-end of each primer and defined in primers' name (Table 1). All restriction enzymes were purchased from *Thermo Fisher Scientific*. I took 0.5 µl of each restriction enzyme (10 U/µl), 500 ng of each DNA sample, and 2 µl of 10x FastDigest green Buffer (total volume of a reaction was 20 µl). The restriction reactions were done at 37 °C for 1 h. The restricted DNA was visualized under UV (260 nm) using the gel electrophoresis method as described in section 2.2.2.2.

#### 2.2.2.9 DNA sequencing

The PCR products were verified by DNA sequencing using primers specific for the pJET1.2/blunt cloning vector: pJET1.2 Forward and pJET1.2 Reverse Sequencing Primers (CloneJET PCR Cloning Kit, *Thermo Fisher Scientific*). For sequencing of the HPV18 2625 bp PCR product synthesized with primers #7 and #8, I also used primer #9 (Table 1), which is complementary to 1765-1785 bp of the HPV18 genome (hereafter, numbering is given

according to the HPV11 and HPV18 reference sequences listed in the PaVE database [1]). DNA sequencing was performed in Estonian Biocenter Core Laboratory, Tartu, Estonia. The correct DNA clones verified by sequencing were chosen for further cloning.

#### 2.2.2.10 Ligation of the obtained PCR products and parental HPV genomes

In order to clone the appropriate DNA fragments from the pJET1.2/blunt vector into HPV-pMC.BESPX parental plasmids, first, I digested all plasmids with the respective restriction enzymes. The restriction was performed as described in section 2.2.2.8. I used 5 µg of parental plasmids, 8 µg of plasmids containing PCR products obtained using primers #1/2, #3/4, and #7/8, and 30 µg of plasmid containing a 137 bp PCR product synthesized with primers #5/6. The fragments were separated using gel electrophoresis, and then the fragments required for subsequent ligation were purified from the agarose gel as described in section 2.2.2.3. The ligation reactions contained 100 ng of parental plasmid, 300 ng of each fragment, 1 µl of 10x reaction Buffer and 1 µl of T4 DNA Ligase. The ligation mixes were incubated at 16 °C overnight and then transformed to competent DH5α. Bacteria were plated on LB-agar containing 25 µg/ml of kanamycin. The plasmid DNA was isolated and verified by the restriction analysis using the relevant restriction enzymes and also by DNA sequencing. The resulted correct constructs were defined as E1-HA HPV11-pMC.BESPX and E1-HA HPV18-pMC.BESPX.

#### 2.2.2.11 Production of minicircles

Replication of the HPV genome containing bacterial DNA is low-efficient and disparate with replication of wtHPV. All viral genomes used in the replication assay of this study were produced using the minicircle DNA technology (Kay et al., 2010).

For producing minicircles, I used *Escherichia coli* strain ZYCY10P3S2T. The genome of this strain encodes specific recombinase ΦC31 and SclI endonuclease under the inducible L-arabinose P<sub>BAD</sub> promotor (ara-BAD operon). I transformed E1-HA HPV11-pMC.BESPX and E1-HA HPV18-pMC.BESPX plasmids into ZYCY10P3S2T strain and plate bacteria on LB-agar plates containing 25 µg/ml of kanamycin. Bacteria were incubated at 37 °C for 18 h. Bacteria from one colony were transferred into 3 ml LB-broth supplemented with 50 µg/ml of kanamycin, incubated at 37 °C, 220 rpm for 8 h, and then transferred into 100 ml of *Difco Terrific* broth containing 25 µg/ml of kanamycin. Bacteria were grown at 37 °C, 220 rpm for 16 h. Then, I added 100 ml of an inducing solution consisting of LB containing 0.4M NaOH and 0.04% L-arabinose and continued incubation at 32 °C, 220 rpm for 8 h. Then, I centrifuged bacteria at RT and 5000 rpm for 10 min and used Plasmid Extraction Mini Kit (*Flavoprep*) to

extract minicircles in a similar way as described in section 2.2.2.6. The resulted E1-HA HPV11 and E1-HA HPV18 minicircles were used in further experiments.

## 2.2.3 Electroporation of the U2OS cell line and isolation of total DNA

### 2.2.3.1 U2OS cell line and used medium

In this study, I used human osteosarcoma cell line U2OS (Geimanen et al., 2011). Cells were grown in 10 cm plates (*Corning Inc.*) in the normal growth medium (NGM) consisting of Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% fetal bovine serum (FBS), penicillin (100 IU/ml), and streptomycin (100 ng/ml) (*Sigma-Aldrich*). Cells were incubated at 37 °C and 5% of CO<sub>2</sub>.

### 2.2.3.2 Transfection of U2OS cells by electroporation

Confluent U2OS cells were split 1:2 approximately 24 h before transfection. On the next day, the medium was removed and the cells were washed with a PBS solution (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na<sub>2</sub>HPO<sub>4</sub>, and 1.47 mM KH<sub>2</sub>PO<sub>4</sub>). Then, the cells were detached using 0.8 ml of a 0.25% Trypsin-EDTA solution and transferred into 6 ml of a fresh NGM. The cells were counted using a Countess Automated Cell Counter (*Thermo Fisher Scientific*). I used approximately  $1.2 \times 10^6$  cells per each electroporation. The required amount of the cells was centrifuged at 20 °C and 1000 rpm for 5 min using Eppendorf Centrifuge 5810R (*Thermo Fisher Scientific*) and then resuspended in 250 µl of the NGM. The cells were mixed with HPV genomes and 50 µg of salmon sperm DNA. The following amounts of the viral genomes were used for electroporation of  $1.2 \times 10^6$  cells: 600 ng of HPV11 and E1-HA HPV11 minicircle DNA and 1600 ng of HPV18 and E1-HA HPV18 minicircle DNA. Suspension of the cells with DNA was transferred in a cuvette for electroporation with a gap size of 4 mm. Electroporation was performed using a Gene Pulser XCell machine (*Bio-Rad Instruments*) at 220 V voltage and 975 µF capacity. After electroporation, I transferred the cells in 4 ml of a fresh NGM and centrifuged them at 20 °C, 1000 rpm for 5 min. Then, I resuspended the cells in 6 ml of the NGM and equally divided them between three wells of 6-well plates (*Corning Inc.*) using 2 ml of cell suspension per one well.

### 2.2.3.3 Extraction of total DNA from U2OS cells

Prior lysis, the cells were washed with 2 ml of PBS and then lysed in 500 µl of Proteinase K buffer (20 mM Tris pH 8.0, 100 mM NaCl, 0.1 g/l mM EDTA, 0.2% SDS). Then, I mixed the lysates with Proteinase K (final concentration 200 µg/ml) using insulin syringe and incubated at 56 °C overnight. Hydrolyzed proteins were removed using the phenol-chloroform extraction

method. I added 500 µl of phenol-chloroform (1:1) mix to each lysate and vortexed them for 15 s. After centrifugation at RT, 13000 rpm for 2 min, I transferred the upper aqueous phase containing DNA to a new 1.5 ml tube. Then, I added 2 volumes of 96% ethanol to each tube to precipitate the DNA. After vigorous mixing with subsequent 30 min of incubation at -20 °C and centrifugation at 4 °C, 15000 rpm for 20 min, supernatants were removed and dried pellets were resuspended in 100 µl of TE (10 mM Tris, 1 mM EDTA pH 8.0) containing 40 ng of RNase A by vortexing. The mixtures were incubated at 37 °C for 1 h. After that, I precipitated DNA as described in section 2.2.2.7 and dissolved the pellets in 40 µl of TE.

## 2.2.4 Southern Blot

### 2.2.4.1 Restriction of extracted total DNA

The obtained total DNA samples containing genomic, mitochondrial, and viral DNA were analyzed using the Southern Blot method. For the restriction, I took 2 µg of samples containing HPV11 and E1-HA HPV11 minicircle DNA and 3 µg of samples containing HPV18 and E1-HA HPV18 minicircle DNA. I used 2 µl of each restriction enzyme and 4 µl of 10x FastDigest green Buffer per one reaction in 40 µl of the total volume. The reactions were performed at 37 °C overnight.

### 2.2.4.2 Transfer of DNA from the gel to a nylon membrane

The DNA fragments obtained after the restriction were separated on a 0.8% agarose gel containing 0.3 µg/ml of ethidium bromide. The gel was run in 1xTAE buffer for 17 h under 25 V in the presence of 8 µg of O'GeneRuler 1kb DNA ladder. In order to denature dsDNA, I incubated the gel in alkali Solution A (0.5M NaOH, 1.5M NaCl) at RT and 100 rpm for 40 min and then washed with distilled water. Secondly, I incubated the gel in Solution B (1M Tris pH 8.0, 1.5M NaCl) at RT and 100 rpm for 30 min. Then, DNA from the gel was transferred to a nylon membrane (*Millipore*) using the upward capillary transfer method in 10x saline-sodium citrate buffer 10x SSC (1.5M NaCl, 150 mM Na<sub>3</sub>C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>) for 18 h.

### 2.2.4.3 Hybridization

After the transfer, DNA was cross-linked to the membrane using UV (*Stratalinker*). After that, to prevent nonspecific binding, I incubated the membranes in 30 ml of a prehybridization solution (16 ml MQ water, 9 ml 20x SSC, 3 ml 50x Denhardt's solution, 1.5 ml 10% SDS, 0.6 ml of denaturated at 100 °C salmon sperm DNA (10 µg/µl)) at 65 °C for 45 min in rolling tubes. For synthesis of hybridization probes, HPV11 and HPV18 minicircle DNAs were initially linearized using HindIII and BglI restriction enzymes, respectively, and then purified from

agarose gel as described in section 2.2.2.3. Hybridization probes were synthesized using 100 ng of linearized and purified HPV11 and HPV18 minicircle DNAs as templates and DecaLabel<sup>TM</sup> DNA Labeling Kit (*Thermo Fisher Scientific*).

First, 10 µl of decanucleotide primers were added to the DNA template in a total volume of 40 µl. The samples were incubated at 100 °C for 10 min and then placed on ice. Next, 3 µl of mix C containing all dNTPs except for dCTP, 4 µl of  $\alpha$ -32P-dCTP isotope (*Hartman Analytics*), and 1 µl of Klenow Fragment lacking 5'-3' exonuclease activity were added, and the samples were incubated at 37 °C for 15 min. Then I added 4 µl of dNTP. The reactions were incubated at 37 °C for additional 15 min and then stopped by adding 1 µl of 0.5M EDTA and denaturation at 100 °C for 10 min. After that, I added HPV11 and HPV18 hybridization probes to the respective membranes incubated in the prehybridization solution. The membranes were incubated at 65 °C overnight. Next day, I washed the membranes with washing solutions pre-heated at 65 °C. First, I washed them twice with solution I (10xSSC, 0.1% SDS) for 5 min, then once with solution II (5xSSC, 0.1% SDS) for 15 min and twice with solution III (0.5xSSC, 0.1% SDS) for 10 min. All washings were performed at 65 °C.

#### 2.2.4.4 Detection of viral DNA signals

After the hybridization, the membranes were exposed to an X-ray film in cassettes at -70 °C for about 5 and 9 h for HPV11 and HPV18, respectively. The X-ray film was developed and fixed using AGFA Developer and Rapid Fixer solutions.

## 2.3 Results

### 2.3.1 Synthesis of fragments of HPV11/HPV18 DNA with the HA-tag encoding sequence

In order to generate new HPV constructs containing the sequence encoding Human influenza hemagglutinin tag (HA-tag, 5'-TACCCATACGATGTTCCAGATTACGCT-3'), I cloned the sequence into E1 ORFs of the HPV genomes. First of all, I used the PCR and primers, listed in Table 1 to synthesize the HA-tag encoding sequence after the 15<sup>th</sup> nucleotide of E1 ORF (numbering is given from the 1<sup>st</sup> ATG codon of E1) in the context of the HPV11 and HPV18 genomes.

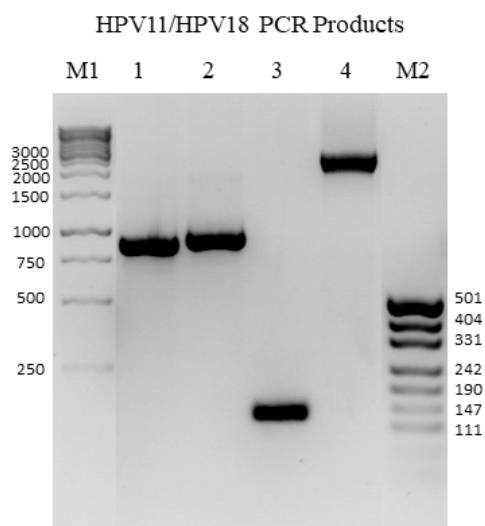
Primers #2 and #6 contain the HA-tag encoding sequence with the NheI restriction site (G↓CTAGC) and also 16 and 17 nucleotides complementary to HPV11 and HPV18 genomic DNAs in positions 823-838 and 912-928, respectively. Primers #3 and #7 are complementary to HPV11 nucleotides 839-855 and HPV18 nucleotides 929-944 and additionally contain the NheI restriction site absent in original HPV11- and HPV18-pMC.BESPX constructs. Primers #1 and #4 are complementary to HPV11 genomic DNA in positions 7883-7905 and 1872-1895, respectively, and have restriction sites BshTI (5'-A↓CCGGT-3') and SwaI (5'-ATTT↓AAAT-3') at their 5' ends. Primers #5 and #8 are complementary to the HPV18 genome in positions 822-843 and 3525-3547, respectively, and contain Bpu1102I (5'-GC↓TNAGC-3') and AatII (5'-GACGT↓C-3') restriction sites at their 5' ends. The obtained DNA amplicons were analyzed using agarose gel electrophoresis (Figure 6).

The expected length of the PCR products obtained using primer pairs #1/2 and #3/4 for HPV11 was 918 bp and 1063 bp, respectively. The expected size of the PCR products synthesized using primer pairs #5/6 and #7/8 for HPV18 was 137 bp and 2625 bp, respectively.

Then, I excised and purified the PCR products from the agarose gel and measured the DNA concentrations. The obtained concentrations of the purified PCR products are listed in Table 4.



**Figure 6**



**Figure 6.** Agarose gel electrophoresis analysis of the fragments of the HPV11 and HPV18 genomes amplified by the PCR. I used DNA ladders O'GeneRuler 1kb (M1) and pUC19 DNA MspI (HpaII) (M2). The obtained PCR products were: (1) 918 bp product of HPV11 primers #1/2, (2) 1063 bp product of HPV11 primers #3/4, (3) 137 bp product of HPV18 primers #5/6, and (4) 2625 bp product of HPV18 primers #5/6. The length of the DNA amplicons matches the expected size.

**Table 4.** Concentrations of the obtained PCR products.

PCR product:	Concentration (ng/μl):
918 bp fragment for HPV11 synthesized with primers #1/2	34
1063 bp fragment for HPV11 synthesized with primers #3/4	53
137 bp fragment for HPV18 synthesized with primers #5/6	66
2625 bp fragment for HPV18 synthesized with primers #7/8	45

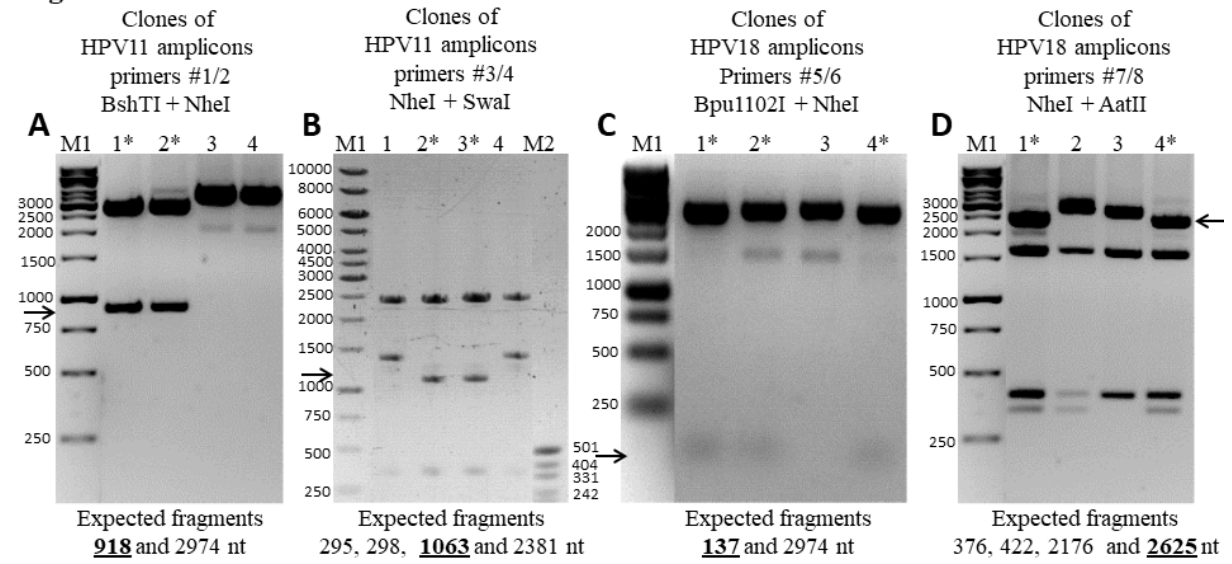
### 2.3.2 Cloning the PCR products into the vector and controlling them with restriction and DNA sequencing

Next, I cloned the obtained DNA fragments into the pJET1.2/blunt cloning vector. In order to verify the presence and size of the cloned DNA fragments, I isolated plasmid DNA from four



independent clones of each variant and performed the restriction analysis using restriction enzymes specific for sites encoded by 5'-ends of primers used for amplification of the fragments (Table 1). Data on the restriction analysis, restriction enzymes used and expected size of the DNA fragments are shown in Figure 7.

**Figure 7**



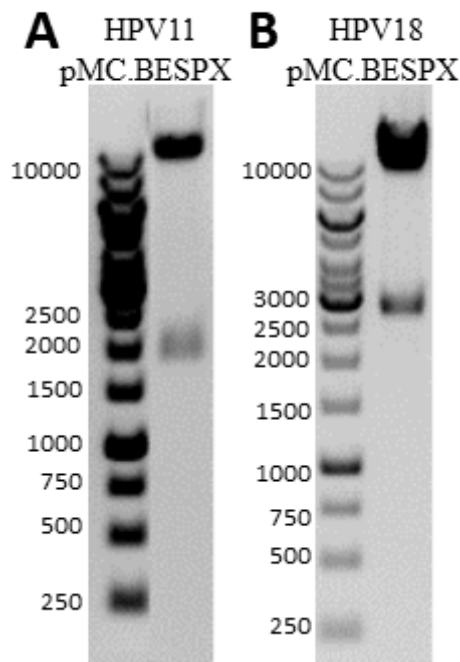
**Figure 7.** Restriction analysis of the PCR products cloned into the pJET1.2/blunt cloning vector. **A:** clones of the HPV11 amplicons synthesized using primers #1/2 were cut using BshTI and NheI restriction enzymes. The expected length of the fragments was 918 and 2974 bp. **B:** clones of the HPV11 amplicons synthesized using primers #3/4 were cut using SwaI and NheI restriction enzymes. Cloning vector pJET1.2/blunt contains two SwaI-specific sites. The expected length of the fragments was 295, 298, 1063, and 2381 bp. **C:** clones of the HPV18 amplicons synthesized using primers #5/6 were cut using Bpu1102I and NheI restriction enzymes. The expected length of the fragments was 137 and 2974 bp. **D:** clones of the HPV18 amplicons synthesized using primers #7/8 were cut using AatII and NheI restriction enzymes. Cloning vector pJET1.2/blunt contains two AatII-specific sites. The expected length of the fragments was 376, 422, 2176, and 2625 bp. The arrows indicate the fragments with the size corresponding to that of the obtained PCR products; the asterisks indicate the clones possessed the expected restriction pattern and selected for sequencing. The numbers corresponding to the sizes of the cloned PCR products are underlined. DNA ladders O'GeneRuler 1kb (M1) and pUC19 DNA Mspl (HpaII) (M2) were used to detect the length of the obtained DNA fragments. The restriction pattern of clones 1(A), 2(A), 2(B), 3(B), 1(C), 2(C), 4(C), 1(D), and 4(D) corresponded to the expected one. Therefore, those clones were selected for DNA sequencing. DNA of clones **1(A), 2(B), 4(C), and 4(D)** had no mutations. These clones were chosen for further assembling with the parental HPV11 and HPV18 genomes.

### 2.3.3 Generation of E1-HA HPV11/HPV18 minicircle constructs

In order to clone the obtained correct DNA fragments into the parental HPV-pMC.BESPX genomes, I restricted HPV11-pMC.BESPX plasmid with BshTI and SmaI and HPV18-pMC.BESPX with Bpu1102I and AatII. The length of the obtained fragments was 1940 bp and 10075 bp for HPV11-pMC.BESPX and 2723 bp and 9251 bp for HPV18-pMC.BESPX. The data are shown in Figure 8.

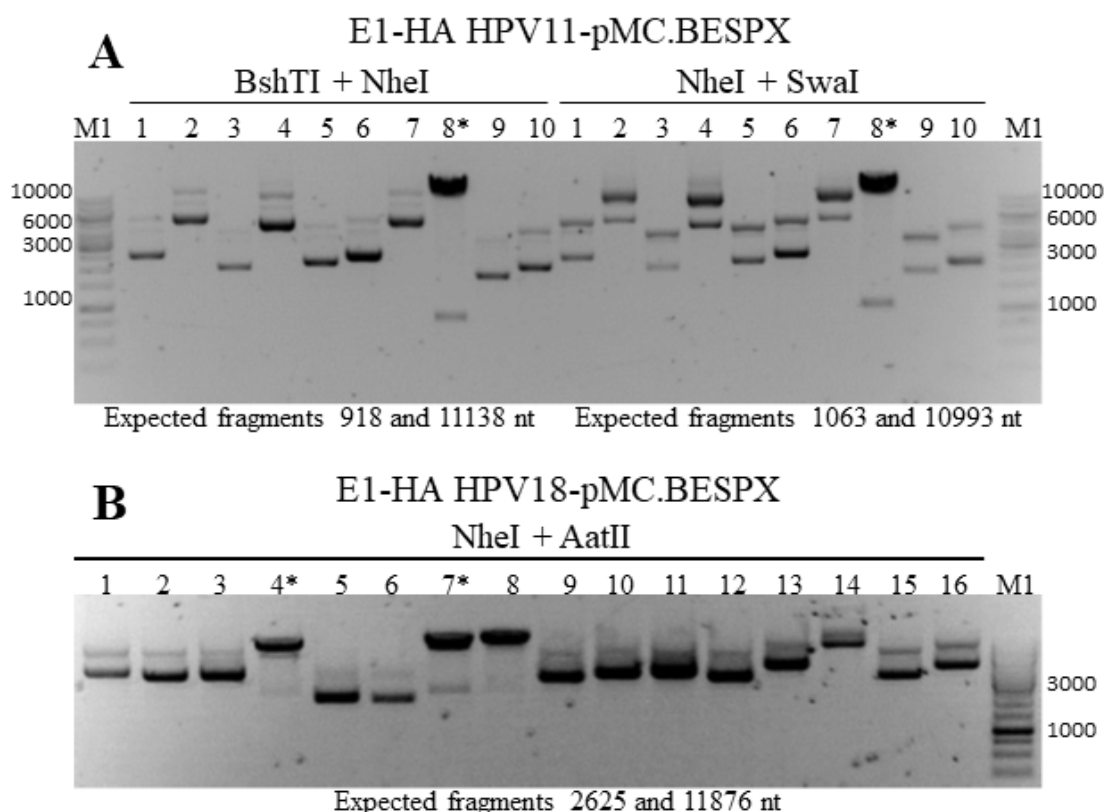
Then, the purified HPV11-pMC.BESPX 10075 bp fragment was ligated with the purified DNA fragments of clones 1(A) (918 bp) and 2(B) (1063 bp). On the other hand, the purified HPV18-pMC.BESPX 9251 bp fragment was ligated with purified DNA fragments of clones 4(C) (137 bp) and 4(D) (2625 bp). The ligation mixtures were transformed into DH5 $\alpha$  bacteria, and plasmid DNA of 10 independent clones of supposed E1-HA HPV11-pMC.BESPX and 16 clones of supposed E1-HA HPV18-pMC.BESPX was purified. After that, the obtained constructs were subjected to the restriction analysis. The data are shown in Figure 9.

**Figure 8**



**Figure 8.** Restriction of the parental HPV-pMC.BESPX genomes. I used O'GeneRuler 1kb as a marker. **A:** restriction of HPV11-pMC.BESPX with BshTI and SmaI restriction enzymes. **B:** restriction of HPV18-pMC.BESPX with Bpu1102I and AatII restriction enzymes. In both cases, the upper bands were excised from the agarose gel and purified.

**Figure 9**

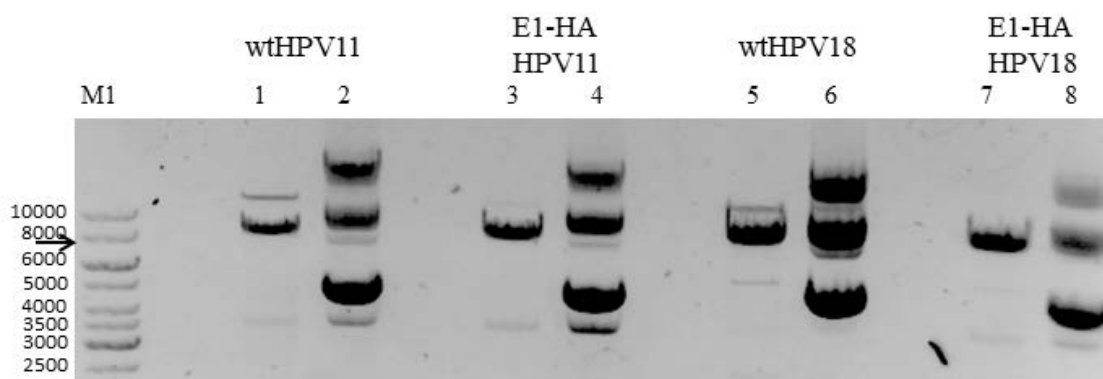


**Figure 9.** Restriction analysis of several clones of the presumed E1-HA HPV11-pMC.BESPX and E1-HA HPV18-pMC.BESPX constructs. **A:** DNA of the supposed E1-HA HPV11-pMC.BESPX clones was digested with BshTI/NheI or NheI/SwaI restriction enzymes. The expected size of the fragments was 918/11138 bp or 1063/10993 bp, respectively. **B:** DNA of the supposed E1-HA HPV18-pMC.BESPX clones was digested with NheI/AatII restriction enzymes. The expected size of the fragments was 2625 bp and 11876 bp. The clones demonstrated the expected restriction pattern are indicated by asterisks. M1: O'GeneRuler 1kb DNA ladder.

Clones **8(A)**, **4(B)**, and **7(B)** were selected for DNA sequencing. Clones **8(A)** and **4(B)** were correct. These clones named E1-HA HPV11-pMC.BESPX and E1-HA HPV18-pMC.BESPX, respectively, were used further for production of the respective minicircle DNAs referred later as E1-HA HPV11 and E1-HA HPV18. During the production of minicircles, a bacterial part of approximately 3 kbp was deleted from the E1-HA HPV11-pMC.BESPX and E1-HA HPV18-pMC.BESPX plasmids and the size of the obtained minicircle DNAs was approximately 8 kbp. Obtained minicircles E1-HA HPV11 and E1-HA HPV18 were compared with the respective wtHPV minicircles using agarose gel electrophoresis. HPV DNA was either linearized or used

in its native form (Figure 10). The conformations and size of the obtained E1-HA HPV minicircles match those of the respective wtHPV genomes.

**Figure 10**



**Figure 10.** Analysis of the conformations and size of the wtHPV and E1-HA HPV minicircles. I used DNA ladder O'GeneRuler 1kb (M1). The HPV11 minicircles were linearized using HindIII restriction enzyme (1) and (3); HPV18 minicircles were linearized using BglI restriction enzyme (5) and (7). Noncut HPV11, E1-HA HPV11, HPV18, and E1-HA HPV18 minicircle DNAs are shown in lanes 2, 4, 6, and 8, respectively.

### 2.3.4 Comparison of the replication efficiencies of wtHPV11 with E1-HA HPV11 and wtHPV18 with E1-HA HPV18 minicircles

The obtained E1-HA HPV11 and E1-HA HPV18 minicircle plasmids were used in the replication assay. I transfected U2OS cells with E1-HA HPV11, wtHPV11, E1-HA HPV18, and wtHPV18 minicircles and incubated cells for 3, 4, and 5 days. After the incubation period, I lysed the cells and extracted the total DNA from the samples. The DNA concentrations are listed in Table 5.

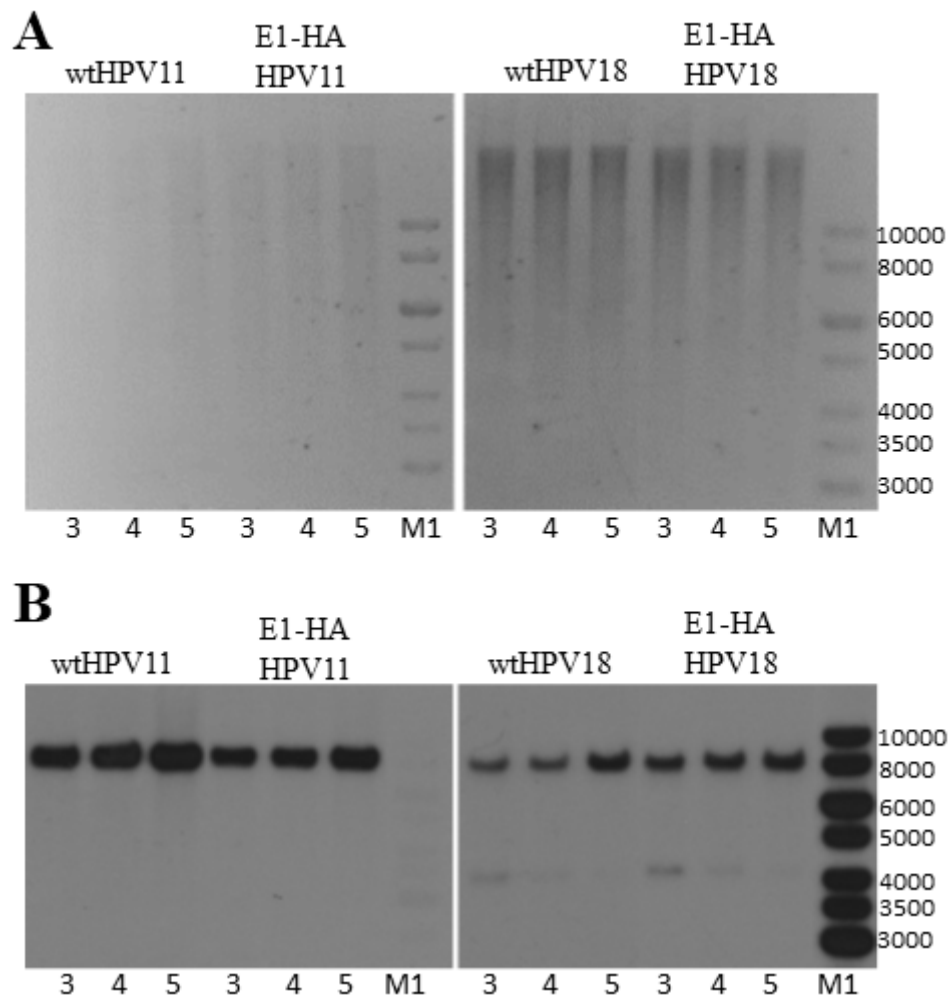
**Table 5.** Concentrations (ng/μl) of the total DNA extracted from U2OS cells transfected with HPV11, E1-HA HPV11, HPV18, and E1-HA HPV18 minicircles.

	3 days:	4 days:	5 days:
wtHPV11	62	215	520
E1-HA HPV11	81	233	575
wtHPV18	223	139	590
E1-HA HPV18	209	272	490

I restricted 2 µg of the total DNA of the samples containing HPV11 and E1-HA HPV11 DNA using restriction enzymes HindIII (10 U/µl) and DpnI (10 U/µl). HindIII linearizes HPV11 DNA by cutting palindromic sequence 5'-A↓AGCTT-3'. DpnI cuts palindromic sequence 5'-Gm6A↓TC-3' in methylated by bacteria DNA, which was used initially for transfection. Treatment with DpnI allows to distinguish HPV DNA replicated in U2OS cells and input DNA used for transfection. For restriction of 3 µg of the total DNA samples containing HPV18 and E1-HA HPV18 minicircle DNA, I used restriction enzymes BglI (10 U/µl) and DpnI. BglI linearizes HPV18 DNA in palindromic sequence 5'-GCCNNNN↓NGGC-3'. The total DNA was separated on the agarose gel and visualized under UV (260 nm) (Figure 11A).

Next, I transferred DNA from the gel to a nylon transfer membrane and hybridized it with radioactively labelled linearized HPV11 and HPV18 minicircle DNA. The data are shown in Figure 11B. The intensity of the radioactive signal at approximately 8 kb represents the level of viral DNA replicated in U2OS cells. The bands detected at approximately 4 kb on the right panel of Figure 11B are the fragments of DpnI-digested input DNA. The Southern blot analysis shows that the replication of E1-HA HPV11 and E1-HA HPV18 minicircles is similar to the respective wild-type genomes in U2OS cells indicating that presence of HA-tag does not impair the E1 biological activity. Therefore, E1-HA HPV11 and E1-HA HPV18 constructs can be used in further studies of E1 proteins and E1-dependent processes.

**Figure 11**



**Figure 11.** Comparison of the replication intensities of wtHPV11/E1-HA HPV11 and wtHPV18/E1-HA HPV18 minicircles in U2OS cells. **A:** total DNA containing wtHPV11, E1-HA HPV11, wtHPV18, and E1-HA HPV18 was visualized under UV. **B:** radioactive signal showing replication of wtHPV11, E1-HA HPV11, wtHPV18, and E1-HA HPV18 minicircles after 3, 4, and 5 days after transfection of U2OS cells.

## 2.4 Discussion

About 65-100% of sexually active people have HPV infection during their life. Most of them are asymptomatic, but sometimes HPV infection may lead to different lesion progressions (Bucchi et al., 2016). About 98% of cervical carcinomas are associated with high-risk mucosal HPV infections (Winters et al., 2006). The study of the HPV infection cycle is essential for prevention and treatment of cancerous diseases.

Since E1 protein is absolutely required for viral replication and inhibition of E1 activities can result in inhibiting viral infection, deeper knowledge of E1 functioning allows to work out novel strategies for manipulating E1 activities and developing new antiviral drugs targeting E1 protein. For development of E1-specific antiviral drugs, it is needed to detect and purify biologically active E1 protein from cells, where HPV is replicating. Detection of endogenous E1 protein is very complicated because of its extremely low levels in host cells and lack of good and universal E1-specific antibodies. Besides, overall similarity of E1 proteins derived from different HPVs is not high, and therefore a unique E1 antibody is needed for each HPV, which makes the research very expensive.

We decided to introduce the universal epitope (HA-tag) into genomes of two different HPV types and compare their and wtHPV replication levels. As a result, the replication patterns of the new HPV genomes were similar to those of wtHPVs in the U2OS cell line suggesting that compared with wtE1, HA-tagged E1 is similarly expressed and has a similar biological activity. The HA-tag is a peptide derived from Human influenza virus. It can be used as an epitope for many commercially available HA-tag specific antibodies. The amino acid sequence of the HA-tag is YPYDVPDYA. Cloning of the HA-tag encoding sequence into E1 ORF of HPV11 and HPV18 provides various possibilities of using E1-HA HPV11/HPV18 constructs in studies on the functions of E1 protein. We can detect the levels and possible post-translational modifications of biologically active HA-tagged E1 protein at different stages of viral replication using specific HA-tag antibodies. Also, we can detect and analyze E1 binding proteins using the co-immunoprecipitation technique. The HA-tag can be cloned into E1 ORFs of other HPV types, generating universal possibility for detecting E1 proteins in different replicating viruses using HA-tag antibodies.

# CONCLUSION

Encoded by HPVs the E1 protein is one of the best targets for the development of new antiviral drugs. The goal of this study is to generate new HPV11 and HPV18 constructs by cloning the HA-tag encoding sequence in the E1 ORF with subsequent testing of their replication ability in human osteosarcoma cell line U2OS. These constructs will help in studies of E1 protein-associated processes, since HA-tag can be used as an epitope for many commercially available HA-tag specific antibodies. The high-risk oncogenic HPV18 and low-risk HPV11 used in the present study belong to mucosal types of Alphapapillomavirus genus. The HA-tag encoding sequence was cloned in their E1 ORFs after the 15<sup>th</sup> nucleotide starting from ATG codon.

First of all, the HA-tag encoding sequence was introduced by the PCR technology into HPV11 and HPV18 DNA fragments, and the obtained fragments were cloned into parental HPV genomes. The resulted constructs were amplified in bacteria, and mutated HPV genomes were generated using the minicircle DNA technology (Kay et al., 2010). Among 10 analyzed clones, only one clone of E1-HA HPV11 was ligated correctly; in the case of E1-HA HPV18, the success rate was two correct clones out of 16 analyzed. Finally, I had fully workable HPV11 and HPV18 E1-HA constructs.

An extrinsical sequence cloned into protein can change its structure and destroy its biological activity. E1 helicase is essential for HPV replication; so I controlled E1-HA HPV11 and E1-HA HPV18 replication ability in U2OS cells and compared their replication intensity with wtHPV11 and wtHPV18 using the Southern Blot method. Replication of different types of HPVs in the U2OS cell line varies greatly, and some of them like HPV11 replicate more intensively than others like HPV18. During the analysis of the replication status of HPVs in the U2OS cell line, I found that the replication of the mutated and wtHPV genomes was similar. The latter indicates that HA-tagged E1 proteins were active.

The obtained E1-HA HPV constructs open a new cost-effective direction in study of E1-associated processes, especially in replication of HPVs. Potentially, the HA-tag encoding sequence can also be cloned into E1 ORFs of other HPVs, and HA-tag specific antibodies can be used in HPV research.



# **Inimese papilloomiviiruse tüüp 11 ja 18 hemaglutiniin märgisega E1 lugemisraami kodeerivate genoomide klonimine ja analüüs**

Viktor Levin

## **RESÜMEE**

Inimese papilloomiviirused (HPV) on laialt levinud populatsioonis ja nad esinevad ka normaalse naha mikrobiotas (Hazard et al., 2007). Nad nakatavad epiteelkoe keratinotsüüte mikrokahjustuse kaudu. Peamine põhjus uurida HPV-d on nende seos vähkkasvajatega, sealhulgas anogenitaal-, pea- ja kaelapiirkonna seotud kasvajatega (Adams et al., 2004). HPV elutsükli parem mõistmine aitab leida uusi meetodeid HPV-seotud haiguseid vältimiseks või ravimiks.

E1 valk on üks parimaid sihtmärke viirusevastaste ravimite arendamiseks. E1 valgul on erinevaid ülesandeid HPV infektsiooni jooksul, kuid E1 põhiülesanne on osaleda HPV DNA replikatsioonis (Bergvall et al., 2013). E1 ja E2 on ainsad HPV valgud, mis on vajalikud ja piisavad HPV replikatsiooniks ja pärssides nende aktiivsust on hea võimalus HPV infektsiooni inhibeerimiseks (Yang et al., 1993). E1 valgu tuvastamine ja puhastamine on suur probleem ja seetõttu on E1-seotud protsesside uurimine ja E1valgu vastu uute ravimite väljatöötamine ka keeruline.

Minu uuringu eesmärk on luua uued HPV11 ja HPV18 genoomide konstruktid koos HA-märgist kodeerivat järjestusega E1 lugemisraami sees ja nende replikatsiooni võime kontrollimine inimese osteosarkoomi U2OS rakuliinis. Esiteks kloonisin HA-märgist kodeerivad järjestused HPV11 ja HPV18 E1 lugemisraami ja seejärel kloonisin saadud fragmendid HPV genoomidesse. Saadud konstruktid paljundasin bakterites ja sain vastavad mutantsed genoomid pisirõngas DNA tehnoloogia abil (Kay et al., 2010). E1 helikaas on vajalik HPV replikatsiooniks ja seega kontrollisin E1-HA HPV11 ja E1-HA HPV18 replikatsioonivõimet võrreldes metsik-tüüpi HPV11 ja HPV18 replikatsiooniga U2OS rakuliinis. E1-HA HPV konstruktide replikatsioon on võrreldav metsik-tüüpi HPV replikatsiooniga ja see tähendab, et HA-märgisega E1 valk on funktsionaalne.

Saadud E1-HA HPV konstruktid avavad uusi võimalusi E1 valguga seotud protsesside, eriti HPV-de replikatsiooni, uurimiseks. Potentsiaalselt on HA-märgist kodeerivat järjestust ka võimalik kloonida teiste HPV-de E1 lugemisraamide sisse ja HA-märgise vastaseid antikehi on võimalik laialt kasutada HPV uurimiseks. See töö on tehtud Tehnoloogiainstituudi molekulaarse viroloogia rühmas Tartu Ülikoolis.

# REFERENCES

- Adams, A. K., Wise-Draper T. M., Wells, S. I. (2014) Human papillomavirus induced transformation in cervical and head and neck cancers. *Cancers* 6(3):1793–820.
- Aksoy, P., Gottschalk, E. Y., Meneses, P. I. (2017) HPV entry into cells. *Mutation Research - Reviews in Mutation Research* 772:13–22.
- Amador-Molina, A., Hernández-Valencia, J. F., Lamoyi, E., Contreras-Paredes, A., Lizano, M. (2013) Role of innate immunity against human papillomavirus (HPV) infections and effect of adjuvants in promoting specific immune response. *Viruses* 5(11):2624–42.
- Amin, A. A., Titolo, S., Pelletier, A., Fink, D., Cordingley, M. G., Archambault, J. (2000) Identification of domains of the HPV11 E1 protein required for DNA replication in vitro. *Virology* 272(1):137–50.
- Anacker, D., Moody, C. (2012) Generation of organotypic raft cultures from primary human keratinocytes. *Journal of Visualized Experiments* (60):3668.
- Ang, K. K., Harris, J., Wheeler, R., ... Gillison, M. L. (2010) Human papillomavirus and survival of patients with oropharyngeal cancer. *The New England Journal of Medicine* 363(1):24–35.
- Archambault, J., Melendy, T. (2015) Targeting human papillomavirus genome replication for antiviral drug discovery. *Antiviral Therapy* 18(3):271–83.
- Berg M., Stenlund A. (1997) Functional interactions between papillomavirus E1 and E2 proteins. *Journal of Virology* 71(5):3853–63.
- Bergvall, M., Melendy T., Archambault, J. (2013) The E1 Proteins. *Virology* 445(1–2):35–56.
- Blitz, I. L., Laimins, L. A. (1991) The 68-kilodalton E1 protein of bovine papillomavirus is a DNA binding phosphoprotein which associates with the E2 transcriptional activator in vitro. *Journal of Virology* 65(2):649–56.
- Bouvard, V., Storey, A., Pim, D., Banks, L. (1994) Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes. *The EMBO Journal* 13(22):5451–59.
- Bravo, I. G., Alonso, Á. (2004) Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *Journal of Virology* 78(24):13613–26.

- Bravo, I. G., de Sanjosé, S., Gottschling, M. (2010) The clinical importance of understanding the evolution of papillomaviruses. *Trends in Microbiology* 18(10):432–38.
- Bucchi, D., Stracci, F., Buonora, N., Masanotti, G. (2016) Human papillomavirus and gastrointestinal cancer: A review. *World Journal of Gastroenterology* 22(33):7415–30.
- Buck, C. B., Day, P. M., Trus, B. L. (2013) The papillomavirus major capsid protein L1. *Virology* 445(1–2):169–74.
- Bzhalava, D., Guan, P., Franceschi, S., Dillner, J., Clifford, G. (2013) A systematic review of the prevalence of mucosal and cutaneous human papillomavirus types. *Virology* 445(1–2):224–31.
- Chabeda, A., Yanez, R. J. R., Lamprecht, R., Meyers, A. E., Rybicki, E. P., Hitzeroth, I. I. (2018) Therapeutic vaccines for high-risk HPV-associated diseases. *Papillomavirus Research* 5:46–58.
- Chen, G., Stenlund, A. (2001) The E1 initiator recognizes multiple overlapping sites in the papillomavirus origin of DNA replication. *Journal of Virology* 75(1):292–302.
- Chow, L. T., Broker, T. R. (2013) Human Papillomavirus Infections: Warts or Cancer? *Cold Spring Harbor Perspectives in Biology* 5(7): a012997.
- Clertant, P., Seif, I. (1984) A common function for polyoma virus Large-T and papillomavirus E1 proteins? *Nature* 311(5983):276–79.
- Clower, R. V., Fisk, J. C., Melendy, C. (2006) Papillomavirus E1 Protein Binds to and Stimulates Human Topoisomerase I. *Journal of Virology* 80(3):1584–87.
- Cueille, N., Nougarede, R., Mechali, F., Philippe, M., Bonne-Andrea, C. (1998) Functional interaction between the bovine papillomavirus virus type 1 replicative helicase E1 and cyclin E-Cdk2. *Journal of Virology* 72(9):7255–62.
- Davy, C. E., Jackson, D. J., Wang, Q., Raj, K., Masterson, P. J., Fenner, N. F., Southern, S., Cuthill, S., Millar, J. B. A., Doorbar, J. (2002) Identification of a G<sub>2</sub> Arrest Domain in the E1<sup>E4</sup> Protein of Human Papillomavirus Type 16. *Journal of Virology* 76(19):9806–18.
- Davy, C., McIntosh, P., Jackson, D. J., Sorathia, R., Miell, M., Wang, Q., Khan, J., Soneji, Y., Doorbar, J. (2009) A Novel Interaction between the Human Papillomavirus Type 16 E2 and E1<sup>E4</sup> Proteins Leads to Stabilization of E2. *Virology* 394(2):266–75.

- De Geest, K., Turyk, M. E., Hosken, M. I., Hudson, J. B., Laimins, L. A., Wilbanks, G. D. (1993) Growth and differentiation of human papillomavirus type 31b positive human cervical cell lines. *Gynecologic Oncology* 49(3):303–10.
- de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U., zur Hausen, H. (2004) Classification of papillomaviruses. *Virology* 324(1):17–27.
- DiMaio, D., Petti, L. M. (2013) The E5 Proteins. *Virology* 445(1–2):99–114.
- Doorbar, J., Parton, A., Hartley, K., Banks, L., Crook, T., Stanley, M., Crawford, L. (1990) Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. *Virology* 178(1):254–62.
- Doorbar, J. (2013) The E4 protein; structure, function and patterns of expression. *Virology* 445(1–2):80–98.
- Egawa, N., Doorbar, J. (2017) The low-risk papillomaviruses. *Virus Research* 231:119–27.
- Favre, M., Breitburd, F., Croissant, O., Orth, G. (1977) Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses. *Journal of Virology* 21(3):1205–09.
- Fradet-Turcotte, A., Brault, K., Titolo, S., Howley, P. M., Archambault, J. (2009) Characterization of papillomavirus E1 helicase mutants’ defective for interaction with the SUMO-conjugating enzyme Ubc9. *Virology* 395(2):190–201.
- Fradet-Turcotte, A., Moody, C., Laimins, L. A., Archambault, J. (2010) Nuclear export of human papillomavirus type 31 E1 is regulated by Cdk2 phosphorylation and required for viral genome maintenance. *Journal of Virology* 84(22):11747–60.
- Geimanen, J., Isok-Paas, H., Pipitch, R., Salk, K., Laos, T., Orav, M., Reinson, T., Ustav, M. Jr., Ustav, M., Ustav, E. (2011) Development of a Cellular Assay System to Study the Genome Replication of High- and Low-Risk Mucosal and Cutaneous Human Papillomaviruses. *Journal of Virology* 85(7):3315–29.
- Graham, S. V. (2010) Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies. *Future Microbiology* 5(10):1493–1506.
- Graham, S. V. (2016) Human Papillomavirus E2 Protein: Linking Replication, Transcription, and RNA Processing. *Journal of Virology* 90(19):8384–88.
- Grassmann, K., Rapp, B., Maschek, H., Petry, K. U., Iftner, T. (1996) Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type

- 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA. *Journal of Virology* 70(4):2339–49.
- Hancock, G., Hellner, K., Dorrell, L. (2018) Therapeutic HPV vaccines. *Best Practice and Research: Clinical Obstetrics and Gynaecology* 47:59–72.
- Hazard, K., Karlsson, A., Andersson, K., Ekberg, H., Dillner, J., Forslund, O. (2007) Cutaneous Human Papillomaviruses Persist on Healthy Skin. *Journal of Investigative Dermatology* 127(1):116–19.
- Howie, H. L., Katzenellenbogen, R. A., Galloway, D. A. (2009) Papillomavirus E6 proteins. *Virology* 384(2):324–34.
- Kadaja, M., Silla, T., Ustav, E., Ustav, M. (2009) Papillomavirus DNA replication — From initiation to genomic instability. *Virology* 384(2):360–68.
- Kajitani, N., Satsuka, A., Kawate, A., Sakai, H. (2012) Productive lifecycle of human papillomaviruses that depends upon squamous epithelial differentiation. *Frontiers in Microbiology* 3:152.
- Kay, M. A., He C-Y., Chen, Z-Y. (2010) A robust system for production of minicircle DNA vectors. *Nature Biotechnology* 28(12):1287–89.
- Ladner, J., Besson, M-H., Audureau, E., Rodrigues, M., Saba, J. (2016) Experiences and lessons learned from 29 HPV vaccination programs implemented in 19 low and middle-income countries, 2009-2014. *BMC Health Services Research* 16(1):575
- Lee, S. M., Park, J. S., Norwitz, E. R., Koo, J. N., Oh, I. H., Park, J. W., Kim, S. M., Kim, Y. H., Park, C. W., Song, Y. S. (2013) Risk of vertical transmission of human papillomavirus throughout pregnancy: a prospective study. *PLoS ONE* 8(6):4–9.
- Leiding, J. W., Holland, S. M. (2012) Warts and All: HPV in Primary Immunodeficiencies. *The Journal of Allergy and Clinical Immunology* 130(5):1030–48.
- Lentz, M. R., Pak, D., Mohr, I., Botchan, M. R. (1993) The E1 replication protein of bovine papillomavirus type 1 contains an extended nuclear localization signal that includes a p34cdc2 phosphorylation site. *Journal of Virology* 67(3):1414–23.
- Letian, T., Tianyu, Z. (2010) Cellular receptor binding and entry of human papillomavirus. *Virology Journal* 7:2.
- Longworth, M. S., Laimins, L. A. (2004) Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiology and Molecular Biology Reviews* 68(2):362–72.

- Loo, Y-M., Melendy, T. (2004) Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA. *Journal of Virology* 78(4):1605–15.
- McBride, A. A. (2008) Replication and Partitioning of Papillomavirus Genomes. *Advances in Virus Research* 72:155–205.
- McBride, A. A. (2013) The Papillomavirus E2 Proteins. *Virology* 445(1–2):57–79.
- Mechali, F., Hsu, C-Y., Castro, A., Lorca, T., Bonne-Andrea, C. (2004) Bovine papillomavirus replicative helicase E1 is a target of the ubiquitin ligase APC. *Journal of Virology* 78(5):2615–19.
- Melendy, T., Sedman, J., Stenlund, A. (1995) Cellular factors required for papillomavirus DNA replication. *Journal of Virology* 69(12):7857–67.
- Mistry, N., Wibom, C., Evander, M. (2008) Cutaneous and mucosal human papillomaviruses differ in net surface charge, potential impact on tropism. *Virology Journal* 5:118.
- Moody, C. A., Fradet-Turcotte, A., Archambault, J., Laimins, L. A. (2007) Human papillomaviruses activate caspases upon epithelial differentiation to induce viral genome amplification. *Proceedings of the National Academy of Sciences of the United States of America* 104(49):19541–46.
- Moody, C. A., Laimins, L. A. (2010) Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer* 10(8):550–60.
- Moody, C. A., Laimins, L. A. (2009) Human Papillomaviruses Activate the ATM DNA Damage Pathway for Viral Genome Amplification upon Differentiation. *PLoS Pathogens* 5(10).
- Oh, S. T., Longworth, M. S., Laimins, L. A. (2004) Roles of the E6 and E7 proteins in the life cycle of low-risk human papillomavirus type 11. *Journal of Virology* 78(5):2620–26.
- Orav, M., Henno, L., Isok-Paas, H., Geimanen, J., Ustav, M., Ustav, E. (2013) Recombination-dependent oligomerization of human papillomavirus genomes upon transient DNA replication. *Journal of Virology* 87(22):12051–68.
- Orav, M., Geimanen, J., Sepp, E-M., Henno, L., Ustav, E., Ustav, M. (2015) Initial amplification of the HPV18 genome proceeds via two distinct replication mechanisms. *Scientific Reports* 5:1–16.
- Park, P., Copeland, W., Yang, L., Wang, T., Botchan, M. R., Mohr, I. J. (1994) The cellular DNA polymerase alpha-primase is required for papillomavirus DNA replication and associates

- with the viral E1 helicase. *Proceedings of the National Academy of Sciences of the United States of America* 91(18):8700–8704.
- Pett, M., Coleman, N. (2007) Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *The Journal of Pathology* 212(4):356–67.
- Prindle, M. J., Loeb, L. A. (2012) DNA polymerase delta in DNA replication and genome maintenance. *Environmental and Molecular Mutagenesis* 53(9):666–82.
- Reinson, T., Toots, M., Kadaja, M., Pipitch, R., Allik, M., Ustav, E., Ustav, M. (2013) Engagement of the ATR-dependent DNA damage response at the human papillomavirus 18 replication centers during the initial amplification. *Journal of Virology* 87(2):951–64.
- Reusser, N. M., Downing, C., Guidry, J., Tying, S. K. (2015) HPV Carcinomas in Immunocompromised Patients. *Journal of Clinical Medicine* 4(2):260–81.
- Richard, A., Tulasne, D. (2012) Caspase cleavage of viral proteins, another way for viruses to make the best of apoptosis. *Cell Death and Disease* 3(3): e277.
- Sakakibara, N., Mitra, R., McBride, A. A. (2011) The papillomavirus E1 helicase activates a cellular DNA damage response in viral replication foci. *Journal of Virology* 85(17):8981–95.
- Sedman, J., Stenlund, A. (1998) The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *Journal of Virology* 72(8):6893–97.
- Singh, N., Senapati, S., Bose, K. (2016) Insights into the mechanism of human papillomavirus E2-induced procaspase-8 activation and cell death. *Scientific Reports* 6(503):1–11.
- Straub, E., Dreer, M., Fertey, J., Iftner, T., Stubenrauch, F. (2014) The viral E8<sup>E2C</sup> repressor limits productive replication of human papillomavirus 16. *Journal of Virology* 88(2):937–47.
- Thorner, L. K., Lim, D. A., Botchan, M. R. (1993) DNA-binding domain of bovine papillomavirus type 1 E1 helicase: structural and functional aspects. *Journal of Virology* 67(10):6000–14.
- Titolo, S., Welchner, E., White, P. W., Archambault, J. (2003) Characterization of the DNA-binding properties of the origin-binding domain of simian virus 40 large T antigen by fluorescence anisotropy. *Journal of Virology* 77(9):5512–18.
- Ustav, E., Ustav, M., Szymanski, P., Stenlund, A. (1993) The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator. *Proceedings of the National Academy of Sciences of the United States of America* 90(3):898–902.

- Ustav, M. Jr. (2016) Molecular Studies of HPV-18 Genome Segregation and Stable Replication. *Dissertationes Technologiae Universitatis Tartuensis*.
- Võsa, L., Sudakov, A., Remm, M., Ustav, M., Kurg, R. (2012) Identification and analysis of papillomavirus E2 protein binding sites in the human genome. *Journal of Virology* 86(1):348–57.
- Wang, J. W., Roden, R. B. (2013a) L2, the minor capsid protein of papillomavirus. *Virology* 445(1–2):175–86.
- Wang, J. W., Roden, R. B. (2013b) Virus-like particles for the prevention of human papillomavirus-associated malignancies. *Expert review of vaccines* 12(2):129–41.
- Wang, X., Meyers, C., Wang, H. K., Chow, L. T., Zheng, Z. M. (2011) Construction of a full transcription map of human papillomavirus type 18 during productive viral infection. *Journal of Virology* 85(16):8080–92.
- Whelan, F., Stead, J. A., Shkumatov, A. V., Svergun, D. I., Sanders, C. M., Antson, A. A. (2012) A flexible brace maintains the assembly of a hexameric replicative helicase during DNA unwinding. *Nucleic Acids Research* 40(5):2271–83.
- Williams, V. M., Filippova, M., Soto, U., Duerksen-Hughes, P. J. (2011) HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future Virology* 6(1):45–57.
- Wilson, R., Ryan, G. B., Knight, G. L., Laimins, L. A., Roberts, S. (2007) The full-length E1E4 protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression. *Virology* 362(2):453–60.
- Winters, U., Roden, R., Kitchener, H., Stern, P. (2006) Progress in the Development of a Cervical Cancer Vaccine. *Therapeutics and Clinical Risk Management* 2(3):259–69.
- Yang, L., Mohr, I., Fouts, E., Lim, D. A., Nohaile, M., Botchan, M. (1993) The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA helicase. *Proceedings of the National Academy of Sciences of the United States of America* 90(11):5086–90.
- Yim, E-K., Park, Y-S. (2005) The Role of HPV E6 and E7 Oncoproteins in HPV-associated Cervical Carcinogenesis. *Cancer Research and Treatment* 37(6):319–24.
- Zhao, K. N., Hengst, K., Liu, W. J., Liu, Y. H., Liu, X. S., McMillan, N. A., Frazer, I. H. (2000) BPV1 E2 protein enhances packaging of full-length plasmid DNA in BPV1 pseudovirions. *Virology* 272(2):382–93.



# USED WEBSITES

[1] <https://pave.niaid.nih.gov> (27.05.18)

[2] <https://talk.ictvonline.org> (27.05.18)

**Non-exclusive license to reproduce this thesis and make it available to the public**

**I, Viktor Levin**

(date of birth: 24.03.1996)

1. grant the University of Tartu a free permit (non-exclusive license) to my thesis:

**Generation of human papillomavirus type 11 and 18 genomes encoding hemagglutinin epitope-tagged E1 proteins,**

supervised by PhD Alla Piirsoo,

1.1. to reproduce this thesis for the purpose of preservation and making it available to the public, including the purpose of adding it in DSpace digital archives up to the expire of the time limit of the validity of the copyright;

1.2. to make this thesis available to the public via the web environment of the University of Tartu, including DSpace digital archives, up to the expire of the time limit of the validity of the copyright.

2. am aware of the fact that the author keeps the rights mentioned in section 1.

3. confirm that the granting of this non-exclusive license does not infringe the intellectual property rights of other persons or rights arising from the Personal Data Protection Act.

Tartu, 27.05.2018